Integrative Modeling of Emotions in Virtual Agents

Siddiqui, G.F.

2010

Integrative Modeling of Emotions in Virtual Agents

Ghazanfar Farooq Siddiqui

Amsterdam, The Netherlands, 2010
SIKS Dissertation Series No. 2010-39

Thesis Reading Committee:
prof. dr. J.-J. Ch. Meyer (Utrecht University, Utrecht, NL)
prof. dr. M.A. Neerincx (TU Delft, Delft, NL)
dr. M.C.A. Klein (VU University, Amsterdam, NL)
dr. P. Roelofsma (VU University, Amsterdam, NL)

The research reported in this thesis has been funded by the Higher Education Commission of Pakistan (HEC).
Internet: http://www.hec.gov.pk

SIKS Dissertation Series No. 2010-39
The research reported in this thesis has been carried out under the auspices of SIKS, the Dutch Research School for Information and Knowledge Systems.

VRIJE UNIVERSITEIT

Integrative Modeling of Emotions in Virtual Agents

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad Doctor aan
de Vrije Universiteit Amsterdam,
op gezag van de rector magnificus
prof.dr. L.M. Bouter,
in het openbaar te verdedigen
ten overstaan van de promotiecommissie
van de faculteit der Exacte Wetenschappen
op dinsdag 28 september 2010 om 13.45 uur
in de aula van de universiteit,
De Boelelaan 1105

door

Ghazanfar Farooq Siddiqui

geboren te Mandi Baha Uddin, Pakistan

promotor:	prof.dr. J. Treur
copromotoren:	dr. T. Bosse
	dr. J.F. Hoorn

Acknowledgements

It was a sunny afternoon of June 6, 2006 when my plane landed at Schiphol airport, Amsterdam. It was my first journey abroad. My Pakistani friends warmly welcomed me at Schiphol. I spent my first night with my friends. The next morning I collected the keys from the DUWO office, and headed towards the university. My friend guided me to reach the AI department of VU Amsterdam. I can still remember the first question Mark asked me: what is your first impression about the Netherlands? My answer was the 'greenness' of the Netherlands. After hearing my answer, he laughed abruptly (giving me the impression that I answered wrongly). Over the last four years I have been tramping a research path, hoping to reach a milestone from which to look back and realize that I had made it. Now I am there. It was truly a great experience. The truth is that I would not have been able to successfully complete this PhD without the encouraging people around me. I was lucky to have colleagues, friends and family to support and help me through this journey and see me through to its end.

First and foremost, I would like to express my gratitude to Allah (SWT), for providing me the blessing and picking me out of the millions of my kind in Pakistan, and making my foreign stay and research possible through His obvious and obscure ways. I thank Him for guiding me through my whole life so far and bestowing on me so much of His bounties which I so little deserved.

I wish to express my profound thanks to my supervisor Jan Treur for providing me the opportunity to work at VU Amsterdam in the ASR group. Both from a logistic and a research point of view, Jan has been very supportive, helpful and patient throughout my stay in Amsterdam. I learnt a lot of professional skills from him, of which scientific research is just a subset. Again I thank Jan Treur for all his personal and professional help.

Next, I would like to thank my daily supervisor Tibor Bosse for his competence, encouragement and endless support during all of my PhD activities. He always has a positive, and helping attitude. I like the way he gives feedback about my research. I am also thankful to Johan F. Hoorn for his guidance as my second co-supervisor. Johan's comments were always concise and very insightful.

Let me also take this opportunity to thank my doctoral reading committee: Prof. dr. John-Jules Meyer, Prof. dr. Mark Neerincx, Dr. Peter Roelofsma and Dr. Michel Klein for their valuable comments and suggestions after reviewing the dissertation.

The office environment plays a key role in one's research life. I therefore also want to thank all of my colleagues of the ASR group at VU Amsterdam who gave valuable feedback throughout my research. Especially, I thank Matthijs with whom I had a very productive collaboration during the last 3 years. At the VU, I really enjoyed the different activities of the ASR group. I will always remember all the outings, bowling events, dinners, Sinterklaas celebrations and especially the Russian and Pakistani night. Further, I would like to acknowledge my colleagues Alexei, Annerieke, Arlette, Azizi, Charlotte, Fiemke, Mark, Natalie, Nataliya, Peter-Paul, Rianne, Robbert-Jan, Umair, Vera, Waqar and Zulfiqar for the wonderful time.

In the Netherlands, I enjoyed the company of (Pakistani / Indian) friends. They are, Asad, Faraz, Irshad, Jamil, Mazhar, Sabir, Sadeeq, Saleem, Shahaab, Shumais, and Suhail. Especially, I will never forget Asim's Biryani at the dinner parties. We had not known each other in Pakistan, but we lived here for the last four years as an extended family. I thank them all.

I owe this thesis work to my family: my very capable parents who did a lot for their kids, my sister Rubina and brothers Qaisar, Ansar, Moazzam and especially Nasir, who guided me in every aspect of my professional life. I like to thank them for their endless love and support during my entire life.

Last but not least, I would like to thank Zomra for being a loving wife and a good friend, and bearing with patience all the hard work and tribulation that came with my research. It's a joy and comfort to share my life with her. In various ways I depended on her: to keep me trouble free and focused on my research and for singly taking care of our children, Ahmad and Haris.

Ghazanfar Farooq Siddiqui
Pakistan, July 2010.

Contents

PART I INTRODUCTION

CHAPTER 1

Introduction

Introduction

1 Background

Over the past years, there has been an increasing interest in the application of intelligent virtual agents in various domains. Intelligent virtual agents (IVAs) are autonomous, graphically embodied agents in a virtual environment that are able to interact intelligently with the environment, other IVAs, or with human users (e.g., [20] and [22]). Typical examples are conversational agents (e.g., [16]), agents in (serious) games (e.g., an instructor in a naval training simulator [10]), and agents in computer-generated virtual stories (e.g., [8]). Recently, much research has been dedicated to developing virtual agents with more realistic graphical representations. However, the affective properties of such agents are usually rather limited, and not very human-like. For example, although many IVAs currently have the ability to somehow show emotions by means of different facial expressions, it is quite difficult for them to show the right emotion at the right moment. One step further, it is even more difficult for them to actually understand and react empathically to the emotional state of other agents. This is in conflict with the requirement of virtual agents to closely mimic human affective behavior. Several studies in Social Sciences have shown that this is an important prerequisite for an agent to increase human involvement in a virtual environment [13]. Therefore, existing systems based on IVAs are not as effective as they could be. Properties that they typically lack are the ability to show emotions (not only in terms of facial expression, but also in terms of behavior), in relation to insight in each other's and humans' cognitive and affective states.

To deal with this problem, some authors propose to increase the affective properties of interactive software agents by using knowledge from psychology and cognitive science as a basis for computational modeling of the cognitive and affective processes involved (e.g., [2], [9] and [15]). Recently, a variety of such computational models have been (and are still being) developed for different aspects of human behavior. Examples include models for reasoning processes [21], visual attention [14], emotion regulation [17], mindreading [18], stress and workload [11], and moods [5]. If such computational models are available in a formal format, this opens the possibility to equip IVAs with them. However, the step from existing computational models (that are mostly used for simulation purposes) to models, that can directly be plugged in to a virtual (3D) environment, in such a way that the IVAs behave according to the cognitive model, is a nontrivial one. In reality, this step involves an iterative process, consisting of, among others, the following sub-tasks: refinement of the computational model, translation to a specific programming environment, and testing and evaluation of the resulting model in the virtual setting.

2 Research Goal

The main research goal for this thesis was to explore how computational models of affect can be integrated within virtual agents. To this end different models have been analyzed, formalized, combined, simulated, and evaluated within applications in health care, business and game context.

3 Research Methodology

To examine the design of virtual agents with appropriate emotional behavior, a generic methodology has been used to analyze models of affective processes. A library of models was developed for aspects that are related to affective behavior, with a focus, among others, on emotion, emotion regulation and the balancing of emotional ambiguities (i.e., involvement-distance trade-offs). Modeling is an important activity in many different areas. A model is a representation of an object, system, or idea in some form other than that of the entity itself [24]. Such a representation describes the most important elements and their relations. Cognitive or affective processes cannot be easily studied directly, as it is not easy to measure things that go on in a human mind. In contrast, one can analyze relatively easily how a model of it behaves.

Computational models are able to imitate what would happen with the processes over time. This feature is called simulation. The result of a simulation is a sequence of states of the elements of the model at subsequent time points. Simulations experiments with computational models can be used to improve the understanding of the processes, and to make predictions. The modeling and simulation cycle consists of four activities:

i. **Conceptualization**: determining the main aspects and their relations;
ii. **Formalization**: specifying the detailed model;
iii. **Simulation**: performing experiments with the model;
iv. **Evaluation**: verifying whether the model behaves as expected, and validating the applicability of the model in practice

If the evaluation reveals that the model has some shortcomings, the process starts again with a new conceptualization phase. The behavior of cognitive and affective processes was studied by conducting a number of simulation experiments under different circumstances and by evaluating the results against patterns and properties expected from the theory the models are based on. Moreover, the internal consistency of the models was verified. For validation, the models have been incorporated in embodied conversational agents that were confronted with real persons in serious game settings or in a clinical setting.

4 Modeling and Implementation Techniques

The reported research has made use of a number of modeling and implementation techniques, which are briefly indicated below.

LEADSTO

Modeling the various aspects involved in the models to be built in an integrated manner poses some challenges. On the one hand, qualitative aspects have to be addressed, such as decisions to regulate one's emotion. On the other hand, quantitative aspects have to be addressed, such as levels of involvement and distance. The modeling approach based on the modeling language LEADSTO [6] fulfils such desiderata. It integrates qualitative, logical aspects such as used in approaches based on temporal logic e.g., [4] with quantitative, numerical aspects such as used in Dynamical Systems Theory (e.g., [3] and [19]). To test whether the behavior of the affect regulation and engagement models shows the same patterns as would be expected from the theory, simulation experiments have been performed using the LEADSTO simulation language [6]. These simulation experiments also have verified the internal consistency of the theory.

In LEADSTO direct temporal dependencies between two state properties in successive states are modeled by executable dynamic properties defined as follows. If α and β are state properties of the form 'conjunction of ground atoms or negations of ground atoms', the notation $\alpha \twoheadrightarrow_{e, f, g, h} \beta$, means:

> *If state property α holds for a certain time interval with duration g,*
> *then after some delay (between e and f)*
> *state property β will hold for a certain time interval of length h.*

Here, atomic state properties can have a qualitative, logical format, such as an expression desire(d), expressing that desire d occurs, or a quantitative, numerical format such as an expression has_value(x, v) which expresses that variable x has value v.

C++

C++ is a general purpose programming language developed by Bjarne Stroustrup, in 1979 at Bell Labs; it was originally named 'C with Classes'. The language C++ is an enhancement of the language C. It is a strictly typed and compiled language. It is also known as a middle level language, because of its low level and high level language features [23]. It was used for simulation purposes, among others, due to its computational efficiency over the LEADSTO simulation environment for more complex multi-agent simulations. However, although the implementation of the models was done in C++, for the model specifications LEADSTO was used.

HTML

HTML stands for Hyper Text Markup Language [26]. It is widely used for files that are posted on the internet and viewed by web browsers. HTML is a relatively simple language. All text, graphics, and design elements of a web page are ‘tagged’ with codes that instruct the web browser how to display the files. Such files are easy to recognize because they contain the file extension such as ‘html’ or ‘htm’. Through HTML one can create structured documents by representing structural semantics for text such as lists, forms, paragraphs, radio buttons, headings etc. It also provides links and quotes to other items. It was used in combination with JAVASCRIPT and scripts provided by the Haptek software to develop web pages for different IVA applications.

JavaScript

JavaScript is an object-oriented, cross platform scripting language [1]. It is a small and lightweight language used to enable programmatic access to objects within both the client application as well as server side applications. It is mainly used in the form of client-side JavaScript. It is used for the development of user interfaces and dynamic websites. JavaScript can be expressed as a dynamic, weakly typed, prototype-based language. JavaScript provides a set of objects, such as Array, Date, and Math, and a set of language elements such as operators, control structures, and statements. It was used in combination with scripting commands provided by the Haptek software [12] (see below), to control the Haptek player within a web browser, in order to develop IVA applications.

PHP

PHP stands for Hypertext Preprocessor. It is a server-side scripting language used for the development of dynamic web pages. PHP is a cross platform and open source software [25]. PHP supports many databases like MySQL, Informix, Oracle, Sybase, Solid, Generic ODBC, etc. PHP files can contain text, HTML tags and scripts and are returned to the browser as plain HTML. PHP originally stood for personal home page. PHP syntax is similar to C and JAVA syntax. PHP was used in combination with MySQL, JAVASCRIPT and HTML to store and retrieve data from databases used in the IVA applications.

Haptek PeoplePutty

People Putty is a 3D character building tool that is used to create personalized characters [12]. People Putty includes a wide selection of character features and functionality, e.g.; it can import graphics files to use as textures, alter facial features and shapes of heads. It can also use 3D accessories like hats and clothes. People putty 3D characters can be hosted in web pages using script provided by the software. One can record sound, and the character will read with its lips synched with speech. The Haptek peoplePutty software [12] was used for creating the virtual agents in different applications. Through this program different faces of the virtual agents were created.

5 Thesis Overview

The thesis is based on a collection of articles. The majority of the chapters are reprints of the refereed papers that have been published elsewhere, or extensions thereof. These articles are exactly the same except for their layout. As a result the overlap between the papers has not been removed, for example concerning the introduction to the modeling approach. Furthermore, the articles can be read separately. All authors are cited in alphabetical order and all can be regarded as having made a comparable contribution to articles presented in the thesis, unless explicitly indicated otherwise. The thesis is composed of six parts.

I. Introduction

This preliminary part describes the topic the thesis is about. After that, the goal of the thesis is formulated. Furthermore, the generic modeling approach, as well as the more specific modeling and implementation techniques that are used throughout the thesis, is described briefly.

II. Modeling Involvement between Agents

Part II describes the extent to which an agent becomes involved with another agent, or stays at a distance from it. These extents depend upon different aspects of the other agent such as ethics (good or bad), aesthetics (beautiful or ugly), epistemics or realism (how realistic or unrealistic the other agent is), similarity (resemblances between the two agents) and affordances the other agent offers as an aid or obstacle for task performance of the agent. Chapter 2 introduces a computational model for involvement-distance trade-offs, based on the existing informal model I-PEFiC (Interactively Perceiving and Experiencing Fictional Characters). The main relations within this model have been formalized as regression equations, using the LEADSTO modeling environment. Furthermore, simulation experiments indicated that the model is adequate for simulating the dynamics of involvement-distance trade-offs and their influence on satisfaction. Chapter 3 addresses an extension of this model, which explains affect-driven interaction with mechanisms for goal-directed behavior. Simulation experiments have been conducted, and it was found that agents preferred affect-driven decision options to rational decision options in situations where choices for low expected utility are irrational. Chapter 4 is an extension of the work described in Chapter 3, which manages that the actions agents undertake have an effect on other agents. The agents change their perceptions and beliefs about other agents if actions are taken.

III. Appraisal, Involvement and Regulation

Part III describes the integration of three affect models: CoMERG (computational model of emotion regulation based on Gross theory), EMA (emotion and adaption) and I-PEFiC^{ADM} (Interactively

Perceiving and Experiencing Fictional Characters, i.e., the model developed in part II). Chapter 5 offers a short comparative expose about the three models concerning the generation and regulation of affect, which each in their own right have been successfully applied in the agent and robot domain. It argues that the three models partly overlap in a consistent manner, and where distinct, they complement one another. This chapter provides an analysis of the theoretical concepts, and presents a blueprint of an integration, as a basis for a more advanced representation of affect simulation in virtual humans. Chapter 6 is an extension of the work described in Chapter 5 in which the integrated model, called Silicon Coppélia, was implemented and simulation experiments were performed to test the behavior of the model. These experiments show that the model can simulate richer agent behaviors than any of the models could have done alone.

IV. Modeling Involvement in Economical Context

Part IV focuses on affective states of a person in economical context. This economical context is considered a large scale multi-agent system consisting of thousands or millions of other agents. The cognitive and affective processes and behavior of a person with respect to these other agents that make up the person's economical context concerns forms of involvement with them. More specifically, in this part it is described how involvement in the form of individual investment decisions depends on a personal risk profile, the state of greed of the person, and the state of the world economy. Chapter 7 explores the differences and similarities between a population-based and (individual) agent-based modeling approach in such an economical context. For this purpose a case study on the interplay between individual greed and the global economy was addressed, and a model was developed inspired by the well-known predator-prey model from the literature, where the predator was metaphorically related to greed and prey to the state of the world economy. Simulation results show that agent-based simulations can be closely approximated by population-based simulations. Chapter 8 presents a different agent-based model of human financial decision making, based on psychological states and characteristics such as greed and a personal risk profile. A number of simulation experiments have been performed, which show the ability of model to mimic investment behavior depending different personality types and the state of the economy.

V. Embodying Emotions in Virtual agents

Part V addresses applications of virtual agents that show emotions. Chapter 9 presents an approach to incorporate emotion regulation as addressed within the psychology literature into virtual characters. To this end, Gross' informal theory of emotion regulation, which was formalized using a dynamical system style modeling approach [7] in previous research, was taken as a basis. The chapter reports how a virtual environment has been created, involving a number of virtual agents, which have been equipped with the formalized model for emotion regulation. Chapter 10 presents a virtual agent that guides a person through the Beck Depression Inventory (BDI) questionnaire, which is used to estimate the severity of a depression. The agent responds empathically to the answers given by the user, by changing its facial expression. This resembles face to face therapy more than existing web-based self-help therapies. Chapter 11 presents an affective agent playing tic-tac-toe with a person. Experimenting with a number of agents under different parameter settings shows that the agent is able to show human-like emotional behavior, and can make decisions based on rationality as well as on affective influences. Chapter 12 addresses a financial investment application. This web-based application is equipped with a virtual agent, which tries to mimic human emotions, for example, related to greed and disappointment when a person makes investment decisions and learns about the returns.

VI. Discussion and Future Work

Finally part VI summarizes the research and discusses its relevance to the main research goal established in the current chapter. Furthermore, this part discusses the areas for future work.

References

1. About JavaScript – MDC, http://developer.mozilla.org/en/docs/About_JavaScript
2. Anderson, J.R., and Lebiere, C., The atomic components of thought. Lawrence Erlbaum Associates, Mahwah, NJ, 1998.
3. Ashby, R. (1960). Design for a Brain. Second Edition. Chapman & Hall, London. First edition 1952.
4. Barringer, H., Fisher, M., Gabbay, D., Owens, R., and Reynolds, M. (1996). The Imperative Future: Principles of Executable Temporal Logic, John Wiley & Sons, 1996.
5. Beck, A.T., Cognitive models of depression. Journal of Cognitive Psychotherapy, An International Quarterly, volume 1, 1987, pp. 5-37.
6. Bosse, T., Jonker, C.M., Meij, L. van der, and Treur, J. (2007). A Language and Environment for Analysis of Dynamics by Simulation. International Journal of Artificial Intelligence Tools. Vol. 16, 2007, pp. 435-464.
7. Bosse, T., Pontier, M., and Treur, J., A Computational Model based on Gross' Emotion Regulation Theory. Cognitive Systems Research Journal, vol. 11, 2010, pp. 211-230.
8. Cavazza, M., Charles, F., Mead, S., Interacting with virtual characters in interactive storytelling. In: Alonso, E., Kudenko, D., and Kazakov, D. (eds.), Adaptive Agents and Multi-Agent Systems. Lecture Notes in Artificial Intelligence, volume 2636, Springer Verlag, 2003, pp. 318–325.
9. Detje, F., Dörner, D., and Schaub, H. (eds.), The Logic of Cognitive Systems: Proceedings of the Fifth International Conference on Cognitive Modeling, ICCM'03. Universitäts-Verlag Bamberg, 2003.

10. Doesburg, W. A. van, Heuvelink, A., and Broek, E. L. van den., TACOP: A cognitive agent for a naval training simulation environment. In M. Pechoucek, D. Steiner, and S. Thompson (eds.), Proceedings of the Industry Track of the Fourth International Joint Conference on Autonomous Agents and Multi-Agent Systems, 2005, pp. 34-41.
11. Endsley, M., Toward a Theory of Situation Awareness in Dynamic Systems. Human Factors, vol. 37, issue 1, 1995, pp. 32–64.
12. Haptek, Inc., http://www.haptek.com
13. Hoorn, J.F., Konijn, E.A., and Van der Veer, G. C., Virtual Reality: Do not augment realism, augment relevance. Upgrade - Human-Computer Interaction: Overcoming Barriers, vol. 4, issue 1, 2003, pp. 18-26.
14. Itti, L. and Koch, C., Computational Modeling of Visual Attention, Nature Reviews Neuroscience, Vol. 2, No. 3, 2001, pp. 194-203.
15. Laird, J.E., Newell, A., and Rosenbloom, P.S., Soar: an architecture for general intelligence. Artificial Intelligence, volume 33, issue 1, 1987, pp. 1-64.
16. Maes, P., Guttman, R. H., and Moukas, A. G., Agents that buy and sell. Communications of the ACM, volume 42, issue 3, 1999, pp. 81-91.
17. Marsella, S., and Gratch, J., Modeling coping behavior in virtual humans: Don't worry, be happy. In Proceedings of Second International Joint Conference on Autonomous Agents and Multiagent Systems, AAMAS'03. ACM Press, 2003, pp. 313-320.
18. Marsella, S.C., Pynadath, D.V., and Read, S.J., PsychSim: Agent-based modeling of social interaction and influence. In: Lovett, M., Schunn, C.D., Lebiere, C., and Munro, P. (eds.), Proceedings of the International Conference on Cognitive Modeling, ICCM 2004, pp. 243-248 Pittsburg, Pensylvania, USA.
19. Port, R.F., and Gelder, T. van (Eds.). (1995). Mind as Motion: Explorations in the Dynamics of Cognition. MIT Press, Cambridge, Mass.
20. Prendinger, H., Lester, J., Ishizuka, M. (eds.), Intelligent Virtual Agents. Proc. of the 8th International Conference on Intelligent Virtual Agents, IVA'08. Springer LNAI, vol. 5208, 2008.
21. Rips, L.J., The psychology of proof: Deductive reasoning in human thinking. Cambridge, MA, MIT Press, 1994.
22. Ruttkay, Z., Kipp, M., Nijholt, A., Vilhjálmsson, H. (eds.), Intelligent Virtual Agents. Proc. of the 9th International Conference on Intelligent Virtual Agents, IVA'09. Springer LNAI, vol. 5773, 2009.
23. Schildt, H. (1998), C++: The Complete Reference. Third edition. Osborme / McGraw-Hill.
24. Shannon, R.F., Systems Simulation, Prentice-Hall, 1975.
25. http://www.php.net
26. http://www.w3.org/TR/html5/introduction.html#introduction

PART II MODELING INVOLVEMENT BETWEEN AGENTS

CHAPTER 2

A Robot's Experience of Another Robot: Simulation

This chapter appeared as Bosse, T., Hoorn, J.F., Pontier, M., and Siddiqui, G.F., A Robot's Experience of Another Robot: Simulation. In: Sloutsky, V., Love, B.C., and McRae, K. (eds.), Proceedings of the 30th International Annual Conference of the Cognitive Science Society, CogSci'08, 2008, pp. 2498-2503.

A Robot's Experience of Another Robot: Simulation

Tibor Bosse[1], Johan F. Hoorn[2], Matthijs Pontier[1,2], and Ghazanfar F. Siddiqui[1,2]

[1]VU University, Department of Artificial Intelligence
De Boelelaan 1081, 1081 HV Amsterdam, The Netherlands
{tbosse, mpr210, ghazanfa}@few.vu.nl
[2]VU University, Center for Advanced Media Research Amsterdam
Buitenveldertselaan 3, 1082 VA Amsterdam, The Netherlands
JF.Hoorn@fsw.vu.nl

Abstract. To develop a robot that is able to recognize and show affective behavior, it should be able to regulate simultaneously occurring tendencies of positive and negative emotions. To achieve this, the current paper introduces a computational model for involvement-distance trade-offs, based on an existing theoretical model. The main mechanisms of this model have been represented as regression equations, using the LEADSTO modeling environment. A number of simulation experiments have been performed, which indicated that the model is adequate for simulating the dynamics of involvement-distance trade-offs and their influence on satisfaction. More specifically, the experiments confirmed the empirical finding that positive features do not exclusively increase involvement.

Keywords: emotions, computational modeling, involvement-distance trade-off, simulation.

1 Introduction

We are a research group on a mission. Our aim is to model aspects of emotion regulation and involvement-distance trade-offs to build a robot that can contribute to the wellbeing of patients in need of psychological support. Once the software is capable of simulating emotion-regulating mechanisms and, where appropriate, can trade involvement for distance (or vice versa), we will develop a prototype virtual therapist that is tested against real patients. This therapist should be capable of recognizing emotional behavior and should respond to that in an emotionally appropriate way. People often find it hard to admit that they are in need of therapy or coaching. Experimenting with a virtual therapist as a support tool for self-diagnosis may help to overcome that barrier more easily.

However, as luring as this far horizon may be, there is quite some groundwork to be done first and this paper is addressing some of the modeling issues involved. In a counterpart paper in this volume [8], we dealt with theoretical matters of humans perceiving a virtual other and the way to formalize this (i.e. fuzzy sets). The idea was that the empirically validated models of human encounters with agents (e.g., [15]) and models of emotion regulation (e.g., [5, 6, 11]) could be integrated and used to do the reverse, have a robot determine its level of engagement with its user and have it choose the appropriate affective response to it.

There are two things we want to do in this paper, leaving aside many of the other important issues (e.g., [8]). First, we want to test the formula of Werners [17] to see whether it represents the trade-off well between involvement (the robot becomes friendly with its user) and distance (the robot stays aloof). This involvement-distance trade-off is the fundamental mechanism in user encounters with agents (e.g., [15]), film characters (e.g., [10]), and photographs of people (e.g., [9]). Although one would expect the notions of "involved" (i.e., the tendency to be friendly) and "at a distance" (i.e., the tendency to stay aloof) to be part of one and the same continuum, the above studies have pointed out that their interplay is in reality more subtle. For example, humans can at the same time find another person attractive but morally repulsive, useful to achieve certain goals but distasteful in his or her manners (e.g., [10]). We want to see whether Werners' fuzzy trade-off works well enough to implement this mechanism in our prototype "emobot" (e.g., [7]).

Second, we want to show that we can simulate the way perceptual and experiential factors feed into the involvement-distance trade-off. A number of factors help establish the involvement-distance trade-off. For the body of empirical work that sustains this view, see [8] . The ethics of the user, for example, whether the user is of good or bad intent (take care of or kill the robot) directly affects how the trade-off develops. Another important factor is the affordances a user offers as an aid or obstacle for task performance. A user that performs a simple task causes less trouble for the machinery of the robot than a demanding user does. Additional factors are the aesthetics of the user (beautiful or ugly), epistemics or realism (realistic or unrealistic), and similarity (user resembles robot or not). The upshot is that not anything positive leads to involvement but can increase distance as well (e.g., [14, 16]). This has to do with the goals of the robot. If it admires the skills of a user but if those same skills mean it is going to be dumped soon, being skilled raises involvement - but for different reasons - distance as well. This redistribution of information runs via the factors of relevance (whether user is important to achieve robot goals such as being needed or timely maintenance), and valence (the expected positive or negative results of interacting with the user). For an overview of the complete model, see [14], Figure 1.

The involvement-distance trade-off that is fed by all these factors is used for affective response selection see [8]. This is the process of (qualitatively) evaluating the emotional significance of events, see [4, 11]. In this paper, we focus explicitly on events related to what (e.g., [6]) calls situation selection, i.e. selecting situations that bring the robot in a more desirable state. Examples are walking away from a person the robot feels uncomfortable with and looking for another conversation partner. Thus, in our model the term satisfaction indicates the level of appreciation that a robot attaches to a certain situation it is in. If this level of current satisfaction is too low, the robot may want to select another situation based on the expected satisfaction in the future situation.

To account for the trade-off between conflicting options (e.g., 'involvement' vs. 'distance') that influence the level of satisfaction, Werners (e.g., [17]) employs the *fuzzy_AND*-operator γ. Following this approach, each feature u in a trade-off has a membership function μ in the fuzzy sets of involvement ($\tilde{I}$) and distance ($\tilde{D}$), which allows the feature to move between the minimum and maximum degree of membership to these sets:

$$\mu_{\widetilde{and}}\ (\mu_{\tilde{I}}(u),\ \mu_{\tilde{D}}(u)) =$$
$$\gamma \cdot \min\{\mu_{\tilde{I}}(u),\ \mu_{\tilde{D}}(u)\} + ((1 - \gamma)(\mu_{\tilde{I}}(u) + \mu_{\tilde{D}}(u)) / n),$$

where $u \in U$, $\gamma \in [0, 1]$, and n is the number of fuzzy sets ($\tilde{I}$ and $\tilde{D}$) for which the mean is calculated[1]. Basically, then, there are two ways to calculate a trade-off, using the ($\gamma \cdot$ min) option or the ($\gamma \cdot$ max):

$\gamma \cdot \max\{\mu_{\tilde{I}}(u), \mu_{\tilde{D}}(u)\} + ((1 - \gamma)(\mu_{\tilde{I}}(u) + \mu_{\tilde{D}}(u)) / n)$

When the robot feels ambivalent about its user ("sympathetic but clumsy"), using either option may lead to quite different results for the level of satisfaction. When the mean of involvement and distance (the part after $(1 - \gamma)$ in the formula) is the same, the ($\gamma \cdot$ min) version favors decision options in which the involvement and distance values are close to each other, i.e., to decision options which involve relatively more doubt. The ($\gamma \cdot$ max) version favors options in which involvement and distance differ more from each other, i.e., options of which the robot feels less ambivalent.

Our hypothesis (H1), then, explores whether our system can simulate counter-intuitive empirical results concerning the influence of features on involvement and distance:

H1: Positive features do not necessarily and exclusively increase involvement. Due to the redistribution of information via relevance and valence, also distance can be increased. (The same mechanism may apply to negative features that partly increase involvement).

2 Simulation Model

To simulate the dynamics of involvement-distance trade-offs and their influence on satisfaction, the theoretical model by Hoorn [8] (2008) was taken as a basis and was implemented in the LEADSTO modeling environment [2]. This environment enables the modeler to represent the most elementary steps of a process in terms of direct temporal dependencies, and features a dedicated simulation tool. In this section, we describe how we represented the basic mechanisms of the model in LEADSTO[2]. First, a number of design decisions had to be made. In particular, we chose to treat what are actually factor levels as single features. We then represented the features of different agents (e.g., their goodness, realism, or beauty) by real numbers between 0 and 1. In addition, the satisfaction an agent had in a particular situation was represented by a real number between 0 and 1. To model the impact of (the perception of) the different features on an agent's satisfaction, (e.g., [8]) proposed to use fuzzy set theory. The idea of fuzzy set theory is that features have membership functions for various sets, which determine to what degree they are member of these sets. Within LEADSTO, this principle can easily be simulated by using regression equations, but only to the extent that we used factor levels for features. The process of extracting factor levels from features would require a more elaborated model, which is beyond the scope of this paper.

[1] In the remainder of this paper, we only address the simple case of n=2.
[2] For details about the model, see part A of the appendix.

Domain

We created a virtual environment that was inhabited by a number of virtual agents. These agents are fans of soccer teams and express this by wearing club clothes of their favorite team. When these agents meet, to a certain degree they are involved with and at a distance towards each other. These tendencies are based on features of the agents, according to the formulas described in this section. Table 1 shows certain variables that were used in these formulas.

Table 1: Variable names and meanings

Variable	Meaning	Range
$\text{Perceived}_{(\text{<Feature>}, \text{A1}, \text{A2})}$	Agent1's perception of a certain feature of Agent2	[0, 1]
$\text{Designed}_{(\text{<Feature>}, \text{A2})}$	Value assigned by 'the designer' to a certain feature of Agent2	[0, 1]
$\text{Bias}_{(\text{A1}, \text{A2}, \text{<Feature>})}$	Bias that Agent1 has about a certain feature of Agent2	[0, 2]
$\text{Skill}_{(\text{A1}, \text{<Language>})}$	Skill of Agent1 in a certain language	[0, 1]
$\text{ExpectedSkill}_{(\text{A1}, \text{A2}, \text{language})}$	Skill Agent1 expects Agent2 to have in a certain language	[0, 1]
$\beta_{\text{factor1} \leftarrow \text{factor2}}$	Regression weight factor2 has for another factor1 for a certain agent	[0, 1]
$\gamma_{\text{inv-dist}}$	Variable that is used to calculate the involvement-distance trade-off	[0, 1]
$\text{Satisfaction}_{(\text{A1}, \text{A2})}$	Indicates to what extent Agent1 is satisfied withAgent2	[0, 1]

Aesthetics

Each agent has a value for '*designed* beautiful / ugly'. This is a value the designer expects to raise in the user, or in another agent, based on general principles of aesthetics. This value could be seen as the mean 'score' an agent receives for its beauty / ugliness from all other agents. This *designed* value has a data-driven influence on how agents perceive the beauty of another agent. The variable *bias* represents the concept-driven influence on how agents perceive other agents' beauty. Depending on its value, a bias may increase or decrease an agent's perception of another agent's beauty. Note that 'another agent' could also be the agent itself. This is represented by the following formulas, given in mathematical format.

$\text{Perceived}_{(\text{Beautiful}, \text{A1}, \text{A2})} = \text{Bias}_{(\text{A1}, \text{A2}, \text{Beautiful})} * \text{Designed}_{(\text{Beautiful}, \text{A2})}$
$\text{Perceived}_{(\text{Ugly}, \text{A1}, \text{A2})} = \text{Bias}_{(\text{A1}, \text{A2}, \text{Ugly})} * \text{Designed}_{(\text{Ugly}, \text{A2})}$

When two agents meet for the first time, they will assign a *perceived value* in the range [0, 1] to each other's beauty and ugliness according to the formulas above. *Bias* in the range [0, 2] is multiplied with the designed value for the feature in the range [0, 1]. If agent A1 has a *bias* of 1 for, for instance, the beauty of agent A2, then A1 does not under- or overestimate the beauty of A2. If the *bias* is bigger than 1, then A1 is relatively positive about the beauty of agent A2. When the resulting value for the perceived feature is bigger than 1, it is set to 1, to prevent the formula from going out of range.

Ethics

In line with soccer tradition, good guys are those who are fan of the club, and bad guys are fan of the opponent. Agents recognize good and bad guys by the club clothes they are wearing. E. g., if agent A1 is a fan of the soccer club Ajax, and agent A2 wears club clothes of Ajax, then A1 will think of A2 as a 'good guy', but if A2 wears club clothes of the rival soccer club Feyenoord, then A1 will think of A2 as a 'bad guy'. This is represented by:

$\text{Perceived}_{(\text{Good, A1, A2})} = \text{Satisfaction}_{(\text{A2, Club})}$
$\text{Perceived}_{(\text{Bad, A1, A2})} = 1 - \text{Satisfaction}_{(\text{A2, Club})}$

These formulas say that when two agents meet for the first time, they will perceive a value in the range [0, 1] for each other's goodness and badness. The perceived goodness is exactly the same value as the level of satisfaction the agent attaches to the club of which the other agent is a fan. The perceived badness is 1 minus the level of satisfaction. If an agent wears neutral clothes, the values of good and bad are assigned according to a variable that reflects the agents' perception of neutral clothes (which is 0.3 for both good and bad in the simulations in this paper).

Epistemics

The first time agents meet, each agent perceives the epistemics (or realism) of itself and other agents, the same way it perceives the aesthetics of itself and other agents:

$\text{Perceived}_{(\text{Realistic, A1, A2})} = \text{Bias}_{(\text{A1, A2,Realistic})} * \text{Designed}_{(\text{Realistic, A2})}$
$\text{Perceived}_{(\text{Unrealistic, A1, A2})} = \text{Bias}_{(\text{A1, A2, Unrealistic})} * \text{Designed}_{(\text{Unrealistic, A2})}$

Affordances

In the simulation model, the languages Urdu, English, and Dutch are used as the affordances to have a conversation about soccer. Each agent has a certain *skill* level for each language. Agents perceive each other's affordances according to the expectations they have about the possibilities to communicate with the other agent, according to the following formulas (where the sum over all languages is taken):

$\text{Perceived}_{(\text{Aid, A1, A2})} = \Sigma(\text{ExpectedSkill}_{(\text{A1, A2, language})} * \text{Skill}_{(\text{A1, language})})$
$\text{Perceived}_{(\text{Obstacle, A1, A2})} =$
$1 - \Sigma(\text{ExpectedSkill}_{(\text{A1, A2, language})} * \text{Skill}_{(\text{A1, language})})$

When two agents meet, they assign a value to each other's affordances (aid and obstacle) in the range [0, 1], using the presuppositions they have about the language skills of the other agent, which is based on outer appearance. E.g., when Agent A2 has a dark skin, in the simulation, Agent A1 will think Agent A2 has good skills in Urdu, average skills in English, and bad skills in Dutch. Because of this, the value of *aid* is calculated by taking the sum of the language skills of Agent A1 multiplied by the language skills A1 expects A2 to have. These expectations of Agent A1 about the language skills of Agent A2 are normalized, and are based on skin color (although politically incorrect, this was convenient for simulation purposes). The perceived value for *obstacle* was 1 minus the calculated value for *aid*.

Similarity

For an agent to perceive its similarity with another agent, it needs to perceive the features of the self. Agents perceive their own features the same way they perceive the aesthetics and epistemics of other agents. Only this time, the *bias* is the bias in self-perception, instead of in the perception of another agent.

$$Perceived_{(Feature,\ A1,\ A1)} = Bias_{(A1,\ A1,\ Feature)} * Designed_{(Feature,\ A1)}$$

Similarity is perceived according to the differences between the agent's perception of its own features (*good*, *bad*, *beautiful*, *ugly*, *realistic* and *unrealistic*) and its perception of the features of the other agent (where the sum over ranges over these six features):

$$Similarity_{(A1,\ A2)} = 1 - (\Sigma(\beta_{sim \leftarrow feature} * abs(Perceived_{(Feature,\ A1,\ A2)} - Perceived_{(Feature,\ A1,\ A1)})))$$

$$Dissimilarity_{(A1,\ A2)} = \Sigma(\beta_{ds \leftarrow feature} * abs(Perceived_{(Feature,\ A1,\ A2)} - Perceived_{(Feature,\ A1,\ A1)}))$$

To calculate the dissimilarity between two agents, the differences between the perceived values for its own features, and those perceived for the other agent are taken. These differences are all added, with a certain (regression) weight β. Similarity is calculated in a similar manner, but with different weights, and 1 was subtracted by the sum of all differences.

Relevance, Valence, Involvement and Distance

The formulas in this paragraph were designed using generalized linear models [12, 13]. Hoorn (2008) shows that the calculated dependent variable (e.g., relevance) is fed by a number of contributing variables. Each contributing variable has a certain main effect on the dependent variable. The size of this main effect is represented by a (regression) weight β the same way as for calculating similarity. When two variables interact, the interaction effect on the dependent variable is calculated by multiplying the product of the values of these two variables with a certain regression weight, which accounts for the interaction effect on the dependent variable. When the interaction is over-additive, the weight will be positive, and when it is under-additive, the weight will be negative.

The formula for the calculation of a variable A that is dependent on the variables B, C, and D, of which C and D interact, would be: $A = \beta_B*B + \beta_C*C + \beta_D*D + \beta_{CD}*C*D$. In this formula, β_B, β_C, and β_D are the (regression) weights for the main effect of variables B, C, and D respectively, and β_{CD} is the (regression) weight for the interaction effect of C and D.

Due to space limitations, the formulas for relevance, valence, involvement and distance are not given completely, but all the effects on the variables are summarized in Table 2.

The formulas are then constructed using the algorithm described above. For theoretical reasons, each variable in Table 2 is in the actual formula split up in two unipolar variables (ethics is split up into *good* and *bad*, valence is split up into *positive valence* and *negative valence*, engagement is split up into *involvement* and *distance*, etc.).

Table 2: Effects of features on relevance, valence, involvement and distance.

Effects on:	**Main effects**	**Interaction effects**
Relevance	Ethics Aesthetics Epistemics Affordances	Ethics x Affordances Ethics x Aesthetics x Epistemics
Valence	Ethics Aesthetics Epistemics Affordances	Ethics x Affordances Ethics x Aesthetics x Epistemics
Engagement	Similarity Relevance Valence	Relevance x Valence

For example, the formula to calculate relevance looks as follows (where ‘Perceived’ has been abbreviated as ‘Perc’).

$$
\begin{aligned}
&\text{Relevance}_{(A1,\ A2)} = \\
&\beta_{rel \leftarrow good} * \text{Perc}_{(Good,\ A1,\ A2)} + \\
&\beta_{rel \leftarrow bad} * \text{Perc}_{(Bad,\ A1,\ A2)} + \\
&\beta_{rel \leftarrow bea} * \text{Perc}_{(Beautiful,\ A1,\ A2)} + \\
&\beta_{rel \leftarrow ugly} * \text{Perc}_{(Ugly,\ A1,\ A2)} + \\
&\beta_{rel \leftarrow real} * \text{Perc}_{(Realistic,\ A1,\ A2)} + \\
&\beta_{rel \leftarrow unr} * \text{Perc}_{(Unrealistic,\ A1,\ A2)} + \\
&\beta_{rel \leftarrow aid} * \text{Perc}_{(Aid,.\ A1,\ A2)} + \\
&\beta_{rel \leftarrow obst} * \text{Perc}_{(Obstacle,\ A1,\ A2)} + \\
&\beta_{rel \leftarrow good\text{-}aid} * \text{Perc}_{(Good,\ A1,\ A2)} * \text{Perc}_{(Aid,\ A1,\ A2)} + \\
&\beta_{rel \leftarrow good\text{-}obst} * \text{Perc}_{(Good,\ A1,\ A2)} * \text{Perc}_{(Obstacle,\ A1,\ A2)} + \\
&\beta_{rel \leftarrow bad\text{-}obst} * \text{Perc}_{(Bad,\ A1,\ A2)} * \text{Perc}_{(Obstacle,\ A1,\ A2)} + \\
&\beta_{rel \leftarrow bad\text{-}aid} * \text{Perc}_{(Bad,\ A1,\ A2)} * \text{Perc}_{(Aid,\ A1,\ A2)} + \\
&\beta_{rel \leftarrow good\text{-}bea\text{-}real} * \text{Perc}_{(Good,\ A1,\ A2)} * \text{Perc}_{(Beautiful,\ A1,\ A2)} * \text{Perc}_{(Realistic,\ A1,\ A2)} + \\
&\beta_{rel \leftarrow good\text{-}bea\text{-}unr} * \text{Perc}_{(Good,\ A1,\ A2)} * \text{Perc}_{(Beautiful,\ A1,\ A2)} * \text{Perc}_{(Unrealistic,\ A1,\ A2)} + \\
&\beta_{rel \leftarrow good\text{-}ugly\text{-}real} * \text{Perc}_{(Good,\ A1,\ A2)} * \text{Perc}_{(Ugly,\ A1,\ A2)} * \text{Perc}_{(Realistic,\ A1,\ A2)} + \\
&\beta_{rel \leftarrow good\text{-}ugly\text{-}unr} * \text{Perc}_{(Good,\ A1,\ A2)} * \text{Perc}_{(Ugly,\ A1,\ A2)} * \text{Perc}_{(Unrealistic,\ A1,\ A2)} + \\
&\beta_{rel \leftarrow bad\text{-}bea\text{-}real} * \text{Perc}_{(Bad,\ A1,\ A2)} * \text{Perc}_{(Beautiful,\ A1,\ A2)} * \text{Perc}_{(Realistic,\ A1,\ A2)} + \\
&\beta_{rel \leftarrow bad\text{-}bea\text{-}unr} * \text{Perc}_{(Bad,\ A1,\ A2)} * \text{Perc}_{(Beautiful,\ A1,\ A2)} * \text{Perc}_{(Unrealistic,\ A1,\ A2)} + \\
&\beta_{rel \leftarrow bad\text{-}ugly\text{-}real} * \text{Perc}_{(Bad,\ A1,\ A2)} * \text{Perc}_{(Ugly,\ A1,\ A2)} * \text{Perc}_{(Realistic,\ A1,\ A2)} + \\
&\beta_{rel \leftarrow bad\text{-}ugly\text{-}unr} * \text{Perc}_{(Bad,\ A1,\ A2)} * \text{Perc}_{(Ugly,\ A1,\ A2)} * \text{Perc}_{(Unrealistic,\ A1,\ A2)}
\end{aligned}
$$

Satisfaction

Within our model, satisfaction is a certain appreciation the agents attach to the possible decisions they can make (about situation selection). They use their expected satisfaction with each option, to decide which option ‘feels’ best for them. Satisfaction is calculated by a trade-off between involvement and distance:

$$
\begin{aligned}
\text{Satisfaction}_{(A1,\ A2)} = \ &\gamma_{inv\text{-}dist} * \max(\text{Involvement}_{(A1,\ A2)}, \text{Distance}_{(A1,\ A2)}) + \\
&(1 - \gamma_{inv\text{-}dist}) * ((\text{Involvement}_{(A1,\ A2)}, \text{Distance}_{(A1,\ A2)}) / n)
\end{aligned}
$$

When there is relatively more involvement, this will lead to a relatively more positive type of approach towards the other agent. Note that a lot of distance also can lead to a high satisfaction, reflecting a desire for a negative approach (“Beat up the soccer opponent”).

The trade-off is calculated using a variant of the fuzzy_AND-operator γ (e.g., [17, 18]). In the simulation experiments, two variants of this formula tested H1. In the min version, a part γ was taken of the minimum of *involvement* and *distance*, and a part (1 -

γ) was taken of the mean of *involvement* and *distance*, as originally proposed by Werners. In the max version, instead of a part γ of the minimum, a part γ of the maximum of *involvement* and *distance* was taken. In this paper, the value for γ is always set to 0.5.

3 Simulation Results

To test our hypothesis, the simulation model introduced in the previous section was used to perform a number of experiments under different parameter settings. In each experiment, three agents were involved, named Harry, Barry, and Gary. The results of these experiments are described below. Due to space limitations, not all parameter settings are shown in this paper[3].

Baseline Condition. To start, an initial experiment was performed, to test whether the model behaves as expected, and as a control condition for the remaining experiments. In this condition, all the agents had a white skin, and wore Ajax clothes. For the chosen regression weights, see part B of the appendix [1] (to give one example, $\beta_{pv \leftarrow good} = 0.3$, which means the regression weight of *good* on *positive valence* is 0.3). Because all language skills for each agent were set to 0, which means they had no language skills at all, they expected not to be able to communicate with each other, which resulted in assigning 1 to *obstacle*, and 0 to *aid* for each other. With the formula for calculating *good* and *bad*, 'good = 1' implies 'bad = 0.' For this reason, the appraisal the agents attached to Ajax were all set to 0.5, which resulted in all agents assigning 0.5 to each other's goodness, as well as their badness. The satisfaction of each agent with Feyenoord and all $\text{Designed}_{(\text{<Feature>})}$ parameters for the agents were set to 0. All *bias* parameters were set to the neutral value of 1 for each agent. For all agents these parameter settings led to an *involvement* of 0.11, a *distance* of 0.25, and a level of *satisfaction* with each other of 0.21. These values are identical, because all agents are identical. This situation functions as a baseline for the following experiments.

Next, in order to test whether our system could simulate counter-intuitive empirical results (H1) concerning the influence of features on involvement and distance, we experimented with changing the values of *aesthetics* and *epistemics*, see Experiment 1 and 2.

Experiment 1: Aesthetics – beautiful vs. ugly. In this experiment, the parameter settings were the same as in the baseline condition, except that in this experiment, Barry was beautiful (beautiful = 1), and Gary was ugly (ugly = 1). Because of this, Harry's involvement towards Barry (0.11→0.18) increased. Surprisingly, also his distance towards Barry (0.25→0.31) increased. Moreover, both Harry's involvement (0.11→0.14) and distance (0.25→0.37) towards Gary increased as well. It is clear that increasing the value for *beautiful* adds relatively more to *involvement*, and increasing the value for *ugly* adds relatively more to *distance*. As beautiful is a positive feature, which would intuitively be expected to *only* increase involvement, and ugly is a negative feature, which would intuitively be expected to *only* increase distance, this corresponds with H1.

[3] For a detailed description of parameter settings, see appendix, part B.

Experiment 2: Epistemics – realistic vs. unrealistic. In this experiment, the parameter settings were the same as in the baseline condition, except that in this experiment, Barry was realistic (realistic = 1), and Gary was unrealistic (unrealistic = 1). Because of this, Harry's involvement towards Barry (0.11→0.14) increased. Surprisingly, also his distance towards Barry (0.25→0.30) increased. Moreover, both Harry's involvement (0.11→0.14) and distance (0.25→0.31) towards Gary increased as well. As a result of the chosen regression weights, these effects were much smaller than the effects of adding beautiful and ugly. Adding to *realistic* adds relatively more to *involvement*, and adding *unrealistic* adds relatively more to *distance*, although this is much less clear than the difference between adding beautiful and ugly. Because realistic is a positive feature, which traditionally is expected to *only* increase involvement, and unrealistic is a negative feature, which in conventional theories would *only* increase distance, this result confirms H1.

Additional Experiments
In addition to the above experiments, we experimented with changing the values of *ethics* and *affordances*. However, within these formulas, 'good = 1' implies 'bad = 0', and 'aid = 1' implies 'obstacle = 0', and vice versa. Because of this, experimenting with these variables was not suitable for testing H1, since it would never be clear whether the changes in *involvement* and *distance* are caused by the increase in *good*, or by the decrease of *bad*, etc. Nevertheless, a number of experiments were performed with these variables, which confirmed that the behavior of that part of the model globally corresponds to the theoretical model (e.g., [8]) [4].

4 Discussion

In this paper, the theoretical model for involvement-distance trade-offs by [8] has been translated into a simulation model in the LEADSTO language [2]. Two main results were established. First, the model turned out to be adequate for simulating the dynamics of involvement-distance trade-offs and their influence on satisfaction. To model the trade-offs, the ($\gamma \cdot$ max) version of Werners' [17] fuzzy_AND operator seemed to provide more plausible results than the ($\gamma \cdot$ min) version, since in ambiguous cases (where an agent experiences a more or less equal amount of involvement and distance simultaneously), it results in a relatively lower value for satisfaction than the ($\gamma \cdot$ min) version. This is explained by the fact that the ($\gamma \cdot$ max) version favors options in which involvement and distance differ much from each other. For example, it favors situations with I=0.2 and D=0.8 over situations with I=D=0.5), whereas for the ($\gamma \cdot$ min) version this is the other way around. Second, and perhaps more surprisingly, it was found that positive features can increase the level of distance, and that negative features can increase involvement. This is explained by the fact that the factor levels do not directly influence involvement and distance, but only indirectly via the factors of similarity, relevance and valence. Although this finding may be counterintuitive, it corresponds to empirical evidence by [14, 16].

As mentioned above, our model was able to exhibit an increase in distance when the only change to the inputs of the model was an increase in a "positive" parameter. This

[4] See part C of the appendix.

success may seem arguable because the model is so complicated that almost any result is "possible." It might have been wiser to identify the simplest model that could exhibit this behavior, so as to identify the necessary/sufficient components that explain this phenomenon.

However true as this may seem theoretically, from an empirical point of view the model's complexity is warranted by years of experimentation (e.g., [10, 15]). More important than complexity yet is the fact that the model excludes phenomena as well. Based on empirical evidence (ibid.), the model asserts that no more than 10 factors are needed to describe the full of human-robot interaction. These studies (ibid.) also show that realism is subordinate to ethical considerations, that no effects are established but through the mediation of relevance and valence, etc.

Yet, however nicely these empirical data were established, the theory as such still suffered from internal inconsistencies and logical blind spots. This is exactly what we repaired in the current paper. As such, simulations cannot count as a test on ecological validity but we did show that we can simulate empirical results in a logically consistent way. In other words, model verification led to theory improvement.

To do so, we had to create a large number of different bias parameters that were set individually and parameters for setting the other individual characteristics, again underscoring the presumed over-complexity of the model. For one thing, the values of these parameters need to be settled empirically and because we could not do so right away, we set them to zero and one - to neutral that is - thereby reducing complexity again. But what we do have now, by making explicit hidden assumptions and creating internal consistency, is a framework to simulate the more complex affective behaviors and have a solid theoretical basis to do further empirical testing.

In future research, the model will be used to test other, more refined hypotheses. For example, it may be explored whether the use of bipolar variables instead of two unipolar variables (e.g., *ethics* instead of *good* and *bad*) provides different results. In addition, the process of extracting factor levels from features may be modeled in more detail, possibly taking more explicit goals of the robot into account. Another direction for future work is to combine the model with an existing computational model for emotion regulation [3]. Whereas the current model focuses on the *elicitation* of emotion, that model addresses the *regulation* of emotion. We expect that both models will smoothly fit together, since the satisfaction that is generated as output of the involvement-distance trade-off can almost directly be used as input to affective situation selection. In that case, current satisfaction is checked for a certain threshold and if it is too low, the robot will evaluate its expected satisfaction in alternative situations. Finally, in a later stage of the project, the model will be validated against empirical data of human affective trade-off processes. As soon as the model has been validated positively, we will start exploring the possibilities to apply it to real humans instead of agents, i.e., to develop a robot that can communicate affectively with humans.

Acknowledgements

The authors are grateful to the anonymous reviewers for their useful comments on an earlier version of this paper.

References

1. Appendix: see at the end tof the chapter (http://www.cs.vu.nl/~tbosse/emobot).
2. Bosse, T., Jonker, C.M., Meij, L. van der, & Treur, J. (2007). A Language and Environment for Analysis of Dynamics by Simulation. International Journal of Artificial Intelligence Tools, vol. 16, no. 3, pp. 435-464.
3. Bosse, T., Pontier, M., & Treur, J. (2007). A dynamical system modeling approach to Gross' model of emotion regulation. In: Lewis, R.L., Polk, T.A., Laird, J.E. (Eds.), Proceedings of the 8th Int. Conference on Cognitive Modeling, ICCM'07. Taylor and Francis, 2007, pp. 187-192.
4. Gratch, J. (2000). Modeling the Interplay Between Emotion and Decision-Making, Proc. of the 9th Conference on Computer Generated Forces and Behavioral Representation, 2000.
5. Gross, J.J. (1998). The Emerging Field of Emotion Regulation: An Integrative Review. Review of General Psychology, vol. 2, no. 3, pp. 271-299.
6. Gross, J.J. (2001). Emotion Regulation in Adulthood: Timing is Everything. Current directions in psychological science, vol. 10, no. 6, pp. 214-219.
7. Hooley, T., Hunking, B., Henry, M., & Inoue, A. (2004). Generation of Emotional Behavior for Non-Player Characters: Development of EmoBot for Quake II, Proceedings of AAAI, San Jose, CA, 2004.
8. Hoorn, J. F. (2008). A Robot's Experience of its User: Theory. In: Sloutsky, V., Love, B.C., and McRae, K. (eds.), Proceedings of the 30th International Annual Conference of the Cognitive Science Society, CogSci'08 (this volume).
9. Konijn, E.A., & Bushman, B.J. (2007). World leaders as movie characters? Perceptions of G.W. Bush, T. Blair, O. Bin Laden, and S. Hussein at the eve of Gulf War II. Media Psychology. In press.
10. Konijn, E.A., & Hoorn, J.F. (2005). Some like it bad. Testing a model for perceiving and experiencing fictional characters. Media Psychology 7(2): 107-144.
11. Marsella, S., & Gratch, J. (2003). Modeling coping behavior in virtual humans: Don't worry, be happy. In Proceedings of Second International Joint Conference on Autonomous Agents and Multiagent Systems, AAMAS'03. ACM Press, pp. 313-320.
12. McCullagh, P., & Nelder, J.A. (1983). Generalized Linear Models (First ed.). London: Chapman and Hall.
13. Nelder, J. A., & Wedderburn, R. W. (1972). Generalized linear models. J. R. Stat. Soc., A135: 370-384.
14. Van Vugt, H.C., Hoorn, J.F., Konijn, E.A., & De Bie Dimitriadou, A. (2006). Affective affordances: Improving interface character engagement through interaction. International Journal of Human-Computer Studies 64(9): 874-888. DOI: 10.1016/j.ijhcs.2006.04.008
15. Van Vugt, H.C., Konijn, E.A., Hoorn, J.F., Eliëns, A., & Keur, I. (2007). Realism is not all! User engagement with task-related interface characters. Interacting with Computers 19(2) 267-280.
16. Van Vugt, H.C., Konijn, E.A., Hoorn, J.F., & Veldhuis, J. (2006). Why fat interface characters are better e-health advisors. Lecture Notes in Artificial Intelligence (LNAI) 4133: 1-13. DOI 10.1007/11821830_1. Available at: http://www.springerlink.com/content/g379221v1w65tu38/fulltext.pdf
17. Werners, B.M. (1988). Aggregation models in mathematical programming. Mitra: 295-319.
18. Zimmermann, H.J. (1994). Fuzzy Set Theory - and its Applications. Boston, MA: Kluwer-Nijhoff.

Appendix A: The complete formulas for Relevance, Valence and Engagement

To increase readability, the names of the beta weights use the following acronyms:

Variable	Acronym
Good	good
Bad	bad
Beautiful	bea
Ugly	ugly
Realistic	real
Unrealistic	unr
Aid	aid
Obstacle	obst
Similarity	sim
Dissimilarity	ds
Relevance	rel
Irrelevance	irr
Positive Valence	pv
Negative Valence	nv
Involvement	inv
Distance	dis

In these formulas, $\beta_{rel\leftarrow good}$ is the regression weight of good on relevance, $\beta_{rel\leftarrow good\text{-}bea\text{-}real}$ is the regression weight of the interaction of good*bad*realistic on relevance, etc.

$Relevance_{(A1,\ A2)} =$

$\beta_{rel\leftarrow good} * Perceived_{(Good,\ A1,\ A2)} +$
$\beta_{rel\leftarrow bad} * Perceived_{(Bad,\ A1,\ A2)} +$
$\beta_{rel\leftarrow bea} * Perceived_{(Beautiful,\ A1,\ A2)} +$
$\beta_{rel\leftarrow ugly} * Perceived_{(Ugly,\ A1,\ A2)} +$
$\beta_{rel\leftarrow real} * Perceived_{(Realistic,\ A1,\ A2)} +$
$\beta_{rel\leftarrow unr} * Perceived_{(Unrealistic,\ A1,\ A2)} +$
$\beta_{rel\leftarrow aid} * Perceived_{(Aid,.\ A1,\ A2)} +$
$\beta_{rel\leftarrow obst} * Perceived_{(Obstacle,\ A1,\ A2)} +$
$\beta_{rel\leftarrow good\text{-}aid} * Perceived_{(Good,\ A1,\ A2)} * Perceived_{(Aid,\ A1,\ A2)} +$
$\beta_{rel\leftarrow good\text{-}obst} * Perceived_{(Good,\ A1,\ A2)} * Perceived_{(Obstacle,\ A1,\ A2)} +$
$\beta_{rel\leftarrow bad\text{-}obst} * Perceived_{(Bad,\ A1,\ A2)} * Perceived_{(Obstacle,\ A1,\ A2)} +$
$\beta_{rel\leftarrow bad\text{-}aid} * Perceived_{(Bad,\ A1,\ A2)} * Perceived_{(Aid,\ A1,\ A2)} +$
$\beta_{rel\leftarrow good\text{-}bea\text{-}real} * Perceived_{(Good,\ A1,\ A2)} * Perceived_{(Beautiful,\ A1,\ A2)} * Perceived_{(Realistic,\ A1,\ A2)} +$
$\beta_{rel\leftarrow good\text{-}bea\text{-}unr} * Perceived_{(Good,\ A1,\ A2)} * Perceived_{(Beautiful,\ A1,\ A2)} * Perceived_{(Unrealistic,\ A1,\ A2)} +$
$\beta_{rel\leftarrow good\text{-}ugly\text{-}real} * Perceived_{(Good,\ A1,\ A2)} * Perceived_{(Ugly,\ A1,\ A2)} * Perceived_{(Realistic,\ A1,\ A2)} +$
$\beta_{rel\leftarrow good\text{-}ugly\text{-}unr} * Perceived_{(Good,\ A1,\ A2)} * Perceived_{(Ugly,\ A1,\ A2)} * Perceived_{(Unrealistic,\ A1,\ A2)} +$
$\beta_{rel\leftarrow bad\text{-}bea\text{-}real} * Perceived_{(Bad,\ A1,\ A2)} * Perceived_{(Beautiful,\ A1,\ A2)} * Perceived_{(Realistic,\ A1,\ A2)} +$
$\beta_{rel\leftarrow bad\text{-}bea\text{-}unr} * Perceived_{(Bad,\ A1,\ A2)} * Perceived_{(Beautiful,\ A1,\ A2)} * Perceived_{(Unrealistic,\ A1,\ A2)} +$
$\beta_{rel\leftarrow bad\text{-}ugly\text{-}real} * Perceived_{(Bad,\ A1,\ A2)} * Perceived_{(Ugly,\ A1,\ A2)} * Perceived_{(Realistic,\ A1,\ A2)} +$
$\beta_{rel\leftarrow bad\text{-}ugly\text{-}unr} * Perceived_{(Bad,\ A1,\ A2)} * Perceived_{(Ugly,\ A1,\ A2)} * Perceived_{(Unrealistic,\ A1,\ A2)}$

$Irrelevance_{(A1,\ A2)} =$

$\beta_{irr\leftarrow good} * Perceived_{(Good,\ A1,\ A2)} +$
$\beta_{iirr\leftarrow bad} * Perceived_{(Bad,\ A1,\ A2)} +$
$\beta_{irr\leftarrow bea} * Perceived_{(Beautiful,\ A1,\ A2)} +$
$\beta_{irr\leftarrow ugly} * Perceived_{(Ugly,\ A1,\ A2)} +$
$\beta_{irr\leftarrow real} * Perceived_{(Realistic,\ A1,\ A2)} +$
$\beta_{irr\leftarrow unr} * Perceived_{(Unrealistic,\ A1,\ A2)} +$
$\beta_{irr\leftarrow aid} * Perceived_{(Aid,.\ A1,\ A2)} +$
$\beta_{irr\leftarrow obst} * Perceived_{(Obstacle,\ A1,\ A2)} +$

$\beta_{irr \leftarrow \text{good-aid}} * \text{Perceived}_{(Good,\ A1,\ A2)} * \text{Perceived}_{(Aid,\ A1,\ A2)} +$
$\beta_{irr \leftarrow \text{good-obst}} * \text{Perceived}_{(Good,\ A1,\ A2)} * \text{Perceived}_{(Obstacle,\ A1,\ A2)} +$
$\beta_{irr \leftarrow \text{bad-obst}} * \text{Perceived}_{(Bad,\ A1,\ A2)} * \text{Perceived}_{(Obstacle,\ A1,\ A2)} +$
$\beta_{irr \leftarrow \text{bad-aid}} * \text{Perceived}_{(Bad,\ A1,\ A2)} * \text{Perceived}_{(Aid,\ A1,\ A2)} +$
$\beta_{irr \leftarrow \text{good-bea-real}} * \text{Perceived}_{(Good,\ A1,\ A2)} * \text{Perceived}_{(Beautiful,\ A1,\ A2)} * \text{Perceived}_{(Realistic,\ A1,\ A2)} +$
$\beta_{irr \leftarrow \text{good-bea-unr}} * \text{Perceived}_{(Good,\ A1,\ A2)} * \text{Perceived}_{(Beautiful,\ A1,\ A2)} * \text{Perceived}_{(Unrealistic,\ A1,\ A2)} +$
$\beta_{irr \leftarrow \text{good-ugly-real}} * \text{Perceived}_{(Good,\ A1,\ A2)} * \text{Perceived}_{(Ugly,\ A1,\ A2)} * \text{Perceived}_{(Realistic,\ A1,\ A2)} +$
$\beta_{irr \leftarrow \text{good-ugly-unr}} * \text{Perceived}_{(Good,\ A1,\ A2)} * \text{Perceived}_{(Ugly,\ A1,\ A2)} * \text{Perceived}_{(Unrealistic,\ A1,\ A2)} +$
$\beta_{irr \leftarrow \text{bad-bea-real}} * \text{Perceived}_{(Bad,\ A1,\ A2)} * \text{Perceived}_{(Beautiful,\ A1,\ A2)} * \text{Perceived}_{(Realistic,\ A1,\ A2)} +$
$\beta_{irr \leftarrow \text{bad-bea-unr}} * \text{Perceived}_{(Bad,\ A1,\ A2)} * \text{Perceived}_{(Beautiful,\ A1,\ A2)} * \text{Perceived}_{(Unrealistic,\ A1,\ A2)} +$
$\beta_{irr \leftarrow \text{bad-ugly-real}} * \text{Perceived}_{(Bad,\ A1,\ A2)} * \text{Perceived}_{(Ugly,\ A1,\ A2)} * \text{Perceived}_{(Realistic,\ A1,\ A2)} +$
$\beta_{irr \leftarrow \text{bad-ugly-unr}} * \text{Perceived}_{(Bad,\ A1,\ A2)} * \text{Perceived}_{(Ugly,\ A1,\ A2)} * \text{Perceived}_{(Unrealistic,\ A1,\ A2}$

$\text{Positive\_Valence}_{(A1,\ A2)} =$
$\beta_{pv \leftarrow \text{good}} * \text{Perceived}_{(Good,\ A1,\ A2)} +$
$\beta_{pv \leftarrow \text{bad}} * \text{Perceived}_{(Bad,\ A1,\ A2)} +$
$\beta_{pv \leftarrow \text{bea}} * \text{Perceived}_{(Beautiful,\ A1,\ A2)} +$
$\beta_{pv \leftarrow \text{ugly}} * \text{Perceived}_{(Ugly,\ A1,\ A2)} +$
$\beta_{pv \leftarrow \text{real}} * \text{Perceived}_{(Realistic,\ A1,\ A2)} +$
$\beta_{pv \leftarrow \text{unr}} * \text{Perceived}_{(Unrealistic,\ A1,\ A2)} +$
$\beta_{pv \leftarrow \text{aid}} * \text{Perceived}_{(Aid,\ A1,\ A2)} +$
$\beta_{pv \leftarrow \text{obst}} * \text{Perceived}_{(Obstacle,\ A1,\ A2)} +$
$\beta_{pv \leftarrow \text{good-aid}} * \text{Perceived}_{(Good,\ A1,\ A2)} * \text{Perceived}_{(Aid,\ A1,\ A2)} +$
$\beta_{pv \leftarrow \text{good-obst}} * \text{Perceived}_{(Good,\ A1,\ A2)} * \text{Perceived}_{(Obstacle,\ A1,\ A2)} +$
$\beta_{pv \leftarrow \text{bad-obst}} * \text{Perceived}_{(Bad,\ A1,\ A2)} * \text{Perceived}_{(Obstacle,\ A1,\ A2)} +$
$\beta_{pv \leftarrow \text{bad-aid}} * \text{Perceived}_{(Bad,\ A1,\ A2)} * \text{Perceived}_{(Aid,\ A1,\ A2)} +$
$\beta_{pv \leftarrow \text{good-bea-real}} * \text{Perceived}_{(Good,\ A1,\ A2)} * \text{Perceived}_{(Beautiful,\ A1,\ A2)} * \text{Perceived}_{(Realistic,\ A1,\ A2)} +$
$\beta_{pv \leftarrow \text{good-bea-unr}} * \text{Perceived}_{(Good,\ A1,\ A2)} * \text{Perceived}_{(Beautiful,\ A1,\ A2)} * \text{Perceived}_{(Unrealistic,\ A1,\ A2)} +$
$\beta_{pv \leftarrow \text{good-ugly-real}} * \text{Perceived}_{(Good,\ A1,\ A2)} * \text{Perceived}_{(Ugly,\ A1,\ A2)} * \text{Perceived}_{(Realistic,\ A1,\ A2)} +$
$\beta_{pv \leftarrow \text{good-ugly-unr}} * \text{Perceived}_{(Good,\ A1,\ A2)} * \text{Perceived}_{(Ugly,\ A1,\ A2)} * \text{Perceived}_{(Unrealistic,\ A1,\ A2)} +$
$\beta_{pv \leftarrow \text{bad-bea-real}} * \text{Perceived}_{(Bad,\ A1,\ A2)} * \text{Perceived}_{(Beautiful,\ A1,\ A2)} * \text{Perceived}_{(Realistic,\ A1,\ A2)} +$
$\beta_{pv \leftarrow \text{bad-bea-unr}} * \text{Perceived}_{(Bad,\ A1,\ A2)} * \text{Perceived}_{(Beautiful,\ A1,\ A2)} * \text{Perceived}_{(Unrealistic,\ A1,\ A2)} +$
$\beta_{pv \leftarrow \text{bad-ugly-real}} * \text{Perceived}_{(Bad,\ A1,\ A2)} * \text{Perceived}_{(Ugly,\ A1,\ A2)} * \text{Perceived}_{(Realistic,\ A1,\ A2)} +$
$\beta_{pv \leftarrow \text{bad-ugly-unr}} * \text{Perceived}_{(Bad,\ A1,\ A2)} * \text{Perceived}_{(Ugly,\ A1,\ A2)} * \text{Perceived}_{(Unrealistic,\ A1,\ A2)}$

$\text{Negative\_Valence}_{(A1,\ A2)} =$
$\beta_{nv \leftarrow \text{good}} * \text{Perceived}_{(Good,\ A1,\ A2)} +$
$\beta_{nv \leftarrow \text{bad}} * \text{Perceived}_{(Bad,\ A1,\ A2)} +$
$\beta_{nv \leftarrow \text{bea}} * \text{Perceived}_{(Beautiful,\ A1,\ A2)} +$
$\beta_{nv \leftarrow \text{ugly}} * \text{Perceived}_{(Ugly,\ A1,\ A2)} +$
$\beta_{nv \leftarrow \text{real}} * \text{Perceived}_{(Realistic,\ A1,\ A2)} +$
$\beta_{nv \leftarrow \text{unr}} * \text{Perceived}_{(Unrealistic,\ A1,\ A2)} +$
$\beta_{nv \leftarrow \text{aid}} * \text{Perceived}_{(Aid,\ A1,\ A2)} +$
$\beta_{nv \leftarrow \text{obst}} * \text{Perceived}_{(Obstacle,\ A1,\ A2)} +$
$\beta_{nv \leftarrow \text{good-aid}} * \text{Perceived}_{(Good,\ A1,\ A2)} * \text{Perceived}_{(Aid,\ A1,\ A2)} +$
$\beta_{nv \leftarrow \text{good-obst}} * \text{Perceived}_{(Good,\ A1,\ A2)} * \text{Perceived}_{(Obstacle,\ A1,\ A2)} +$
$\beta_{nv \leftarrow \text{bad-obst}} * \text{Perceived}_{(Bad,\ A1,\ A2)} * \text{Perceived}_{(Obstacle,\ A1,\ A2)} +$
$\beta_{nv \leftarrow \text{bad-aid}} * \text{Perceived}_{(Bad,\ A1,\ A2)} * \text{Perceived}_{(Aid,\ A1,\ A2)} +$
$\beta_{nv \leftarrow \text{good-bea-real}} * \text{Perceived}_{(Good,\ A1,\ A2)} * \text{Perceived}_{(Beautiful,\ A1,\ A2)} * \text{Perceived}_{(Realistic,\ A1,\ A2)} +$
$\beta_{nv \leftarrow \text{good-bea-unr}} * \text{Perceived}_{(Good,\ A1,\ A2)} * \text{Perceived}_{(Beautiful,\ A1,\ A2)} * \text{Perceived}_{(Unrealistic,\ A1,\ A2)} +$
$\beta_{nv \leftarrow \text{good-ugly-real}} * \text{Perceived}_{(Good,\ A1,\ A2)} * \text{Perceived}_{(Ugly,\ A1,\ A2)} * \text{Perceived}_{(Realistic,\ A1,\ A2)} +$
$\beta_{nv \leftarrow \text{good-ugly-unr}} * \text{Perceived}_{(Good,\ A1,\ A2)} * \text{Perceived}_{(Ugly,\ A1,\ A2)} * \text{Perceived}_{(Unrealistic,\ A1,\ A2)} +$
$\beta_{nv \leftarrow \text{bad-bea-real}} * \text{Perceived}_{(Bad,\ A1,\ A2)} * \text{Perceived}_{(Beautiful,\ A1,\ A2)} * \text{Perceived}_{(Realistic,\ A1,\ A2)} +$
$\beta_{nv \leftarrow \text{bad-bea-unr}} * \text{Perceived}_{(Bad,\ A1,\ A2)} * \text{Perceived}_{(Beautiful,\ A1,\ A2)} * \text{Perceived}_{(Unrealistic,\ A1,\ A2)} +$
$\beta_{nv \leftarrow \text{bad-ugly-real}} * \text{Perceived}_{(Bad,\ A1,\ A2)} * \text{Perceived}_{(Ugly,\ A1,\ A2)} * \text{Perceived}_{(Realistic,\ A1,\ A2)} +$
$\beta_{nv \leftarrow \text{bad-ugly-unr}} * \text{Perceived}_{(Bad,\ A1,\ A2)} * \text{Perceived}_{(Ugly,\ A1,\ A2)} * \text{Perceived}_{(Unrealistic,\ A1,\ A2}$

$$\begin{aligned}
\text{Involvement}_{(A1,\,A2)} = \quad & \\
& \beta_{inv \leftarrow sim} * \text{Similarity}_{(A1,\,A2)} + \\
& \beta_{inv \leftarrow ds} * \text{Dissimilarity}_{(A1,\,A2)} + \\
& \beta_{inv \leftarrow rel} * \text{Relevance}_{(A1,\,A2)} + \\
& \beta_{inv \leftarrow irr} * \text{Irrelevance}_{(A1,\,A2)} + \\
& \beta_{inv \leftarrow pv} * \text{Pos_Valence}_{(A1,\,A2)} + \\
& \beta_{inv \leftarrow nv} * \text{Neg_Valence}_{(A1,\,A2)} + \\
& \beta_{inv \leftarrow rel\text{-}pv} * \text{Relevance}_{(A1,\,A2)} * \text{Pos_Valence}_{(A1,\,A2)} + \\
& \beta_{inv \leftarrow rel\text{-}nv} * \text{Relevance}_{(A1,\,A2)} * \text{Neg_Valence}_{(A1,\,A2)} + \\
& \beta_{inv \leftarrow irr\text{-}pv} * \text{Irrelevance}_{(A1,\,A2)} * \text{Pos_Valence}_{(A1,\,A2)} + \\
& \beta_{inv \leftarrow irr\text{-}nv} * \text{Irrelevance}_{(A1,\,A2)} * \text{Neg_Valence}_{(A1,\,A2)}
\end{aligned}$$

$$\begin{aligned}
\text{Distance}_{(A1,\,A2)} = \quad & \\
& \beta_{dis \leftarrow sim} * \text{Similarity}_{(A1,\,A2)} + \\
& \beta_{dis \leftarrow ds} * \text{Dissimilarity}_{(A1,\,A2)} + \\
& \beta_{dis \leftarrow rel} * \text{Relevance}_{(A1,\,A2)} + \\
& \beta_{dis \leftarrow irr} * \text{Irrelevance}_{(A1,\,A2)} + \\
& \beta_{dis \leftarrow pv} * \text{Pos_Valence}_{(A1,\,A2)} + \\
& \beta_{dis \leftarrow nv} * \text{Neg_Valence}_{(A1,\,A2)} + \\
& \beta_{dis \leftarrow rel\text{-}pv} * \text{Relevance}_{(A1,\,A2)} * \text{Pos_Valence}_{(A1,\,A2)} + \\
& \beta_{dis \leftarrow rel\text{-}pv} * \text{Relevance}_{(A1,\,A2)} * \text{Neg_Valence}_{(A1,\,A2)} + \\
& \beta_{dis \leftarrow rel\text{-}pv} * \text{Irrelevance}_{(A1,\,A2)} * \text{Pos_Valence}_{(A1,\,A2)} + \\
& \beta_{dis \leftarrow irr\text{-}nv} * \text{Irrelevance}_{(A1,\,A2)} * \text{Neg_Valence}_{(A1,\,A2)}
\end{aligned}$$

Appendix B: Parameter settings in the experiments

This appendix contains all parameter settings for the experiments in the paper. These are the parameter settings of the baseline experiment. The variables that are changed in other experiments are mentioned in the paper.

Table 1: All values for the regression weights

Weight of X	**on Y**	**Value**
Good	Similarity	0.30
Bad	Similarity	0.20
Beautiful	Similarity	0.20
Ugly	Similarity	0.10
Realistic	Similarity	0.10
Unrealistic	Similarity	0.10
Good	Dissimilarity	0.20.
Bad	Dissimilarity	0.30
Beautiful	Dissimilarity	0.10
Ugly	Dissimilarity	0.20
Realistic	Dissimilarity	0.10
Unrealistic	Dissimilarity	0.10
Good	Relevance	0.05
Bad	Relevance	0.05
Beautiful	Relevance	0.08
Ugly	Relevance	0.07
Realistic	Relevance	0.05
Unrealistic	Relevance	0.05
Aid	Relevance	0.05
Obstacle	Relevance	0.05
Good * Aid	Relevance	0.05
Good * Obstacle	Relevance	0.08
Bad * Aid	Relevance	0.08
Bad * Obstacle	Relevance	0.07
Good * Beautiful * Realistic	Relevance	0.03
Good * Beautiful * Unrealistic	Relevance	0.03
Good * Ugly * Realistic	Relevance	0.03
Good * Ugly * Unrealistic	Relevance	0.03
Bad * Beautiful * Realistic	Relevance	0.04
Bad * Beautiful * Unrealistic	Relevance	0.03
Bad * Ugly * Realistic	Relevance	0.03
Bad * Ugly * Unrealistic	Relevance	0.03
Good	Irrelevance	0.05
Bad	Irrelevance	0.05
Beautiful	Irrelevance	0.08
Ugly	Irrelevance	0.07
Realistic	Irrelevance	0.05
Unrealistic	Irrelevance	0.05
Aid	Irrelevance	0.05
Obstacle	Irrelevance	0.05
Good * Aid	Irrelevance	0.07
Good * Obstacle	Irrelevance	0.08
Bad * Aid	Irrelevance	0.08
Bad * Obstacle	Irrelevance	0.07
Good * Beautiful * Realistic	Irrelevance	0.03
Good * Beautiful * Unrealistic	Irrelevance	0.03
Good * Ugly * Realistic	Irrelevance	0.03

Good * Ugly * Unrealistic	Irrelevance	0.03
Bad * Beautiful * Realistic	Irrelevance	0.04
Bad * Beautiful * Unrealistic	Irrelevance	0.03
Bad * Ugly * Realistic	Irrelevance	0.03
Bad * Ugly * Unrealistic	Irrelevance	0.03
Good	Positive Valence	0.30
Bad	Positive Valence	0
Beautiful	Positive Valence	0.15
Ugly	Positive Valence	0
Realistic	Positive Valence	0.03
Unrealistic	Positive Valence	0.02
Aid	Positive Valence	0.30
Obstacle	Positive Valence	0
Good * Aid	Positive Valence	0.30
Good * Obstacle	Positive Valence	-0.10
Bad * Aid	Positive Valence	-0.20
Bad * Obstacle	Positive Valence	0
Good * Beautiful * Realistic	Positive Valence	0.10
Good * Beautiful * Unrealistic	Positive Valence	0.05
Good * Ugly * Realistic	Positive Valence	0.10
Good * Ugly * Unrealistic	Positive Valence	0.05
Bad * Beautiful * Realistic	Positive Valence	-0.10
Bad * Beautiful * Unrealistic	Positive Valence	-0.05
Bad * Ugly * Realistic	Positive Valence	0.03
Bad * Ugly * Unrealistic	Positive Valence	0.02
Good	Negative Valence	0.10
Bad	Negative Valence	0.25
Beautiful	Negative Valence	0.05
Ugly	Negative Valence	0.15
Realistic	Negative Valence	0.03
Unrealistic	Negative Valence	0.07
Aid	Negative Valence	0
Obstacle	Negative Valence	0.35
Good * Aid	Negative Valence	-0.05
Good * Obstacle	Negative Valence	-0.10
Bad * Aid	Negative Valence	-0.15
Bad * Obstacle	Negative Valence	0.30
Good * Beautiful * Realistic	Negative Valence	0.10
Good * Beautiful * Unrealistic	Negative Valence	0.05
Good * Ugly * Realistic	Negative Valence	-0.10
Good * Ugly * Unrealistic	Negative Valence	-0.15
Bad * Beautiful * Realistic	Negative Valence	0.15
Bad * Beautiful * Unrealistic	Negative Valence	0.10
Bad * Ugly * Realistic	Negative Valence	-0.15
Similarity	Involvement	0.15
Dissimilarity	Involvement	0.05
Relevance	Involvement	0.35
Irrelevance	Involvement	-0.1
Positive Valence	Involvement	0.2
Negative Valence	Involvement	0.05
Relevance * Positive Valence	Involvement	0.35
Relevance * Negative Valence	Involvement	0
Irrelevance * Positive Valence	Involvement	0.05
Irrelevance * Negative Valence	Involvement	-0.1
Similarity	Distance	0.05
Dissimilarity	Distance	0.15
Relevance	Distance	0.35

Irrelevance	Distance	-0.1
Positive Valence	Distance	0
Negative Valence	Distance	0.25
Relevance * Positive Valence	Distance	0
Relevance * Negative Valence	Distance	0.4
Irrelevance * Positive Valence	Distance	-0.1
Irrelevance * Negative Valence	Distance	0

Table 2: Levels of satisfaction the agents attach to soccer clubs

Satisfaction of	With	Value
Harry	Ajax	0.5
Harry	Feyenoord	0
Barry	Ajax	0.5
Barry	Feyenoord	0
Gary	Ajax	0.5
Gary	Feyenoord	0

Table 3: The goodness and badness agents attach to other agents wearing neutral clothes

Agent	Clothes	Feature	Value
Harry	Neutral clothes	Good	0.3
Harry	Neutral clothes	Bad	0.3
Barry	Neutral clothes	Good	0.3
Barry	Neutral clothes	Bad	0.3
Gary	Neutral clothes	Good	0.3
Gary	Neutral clothes	Bad	0.3

Table 4: Designed values for the features of each agent

Agent	Feature	Value
Harry	Beautiful	0
Harry	Ugly	0
Harry	Good	0
Harry	Bad	0
Harry	Realistic	0
Harry	Unrealistic	0
Barry	Beautiful	0
Barry	Ugly	0
Barry	Good	0
Barry	Bad	0
Barry	Realistic	0
Barry	Unrealistic	0
Gary	Beautiful	0
Gary	Ugly	0
Gary	Good	0
Gary	Bad	0
Gary	Realistic	0
Gary	Unrealistic	0

Table 5: Biases the agents have in perceiving features of their selves and others

Bias of Agent	For perceiving	Of Agent	Value
Harry	Beautiful	Harry	1
Harry	Beautiful	Barry	1
Harry	Beautiful	Gary	1
Harry	Ugly	Harry	1
Harry	Ugly	Barry	1
Harry	Ugly	Gary	1
Harry	Realistic	Harry	1
Harry	Realistic	Barry	1
Harry	Realistic	Gary	1
Harry	Unrealistic	Harry	1
Harry	Unrealistic	Barry	1
Harry	Unrealistic	Gary	1
Harry	Good	Harry	1
Harry	Good	Barry	1
Harry	Good	Gary	1
Harry	Bad	Harry	1
Harry	Bad	Barry	1
Harry	Bad	Gary	1
Barry	Beautiful	Harry	1
Barry	Beautiful	Barry	1
Barry	Beautiful	Gary	1
Barry	Ugly	Harry	1
Barry	Ugly	Barry	1
Barry	Ugly	Gary	1
Barry	Realistic	Harry	1
Barry	Realistic	Barry	1
Barry	Realistic	Gary	1
Barry	Unrealistic	Harry	1
Barry	Unrealistic	Barry	1
Barry	Unrealistic	Gary	1
Barry	Good	Harry	1
Barry	Good	Barry	1
Barry	Good	Gary	1
Barry	Bad	Harry	1
Barry	Bad	Barry	1
Barry	Bad	Gary	1
Gary	Beautiful	Harry	1
Gary	Beautiful	Barry	1
Gary	Beautiful	Gary	1
Gary	Ugly	Harry	1
Gary	Ugly	Barry	1
Gary	Ugly	Gary	1
Gary	Realistic	Harry	1
Gary	Realistic	Barry	1
Gary	Realistic	Gary	1
Gary	Unrealistic	Harry	1
Gary	Unrealistic	Barry	1
Gary	Unrealistic	Gary	1
Gary	Good	Harry	1
Gary	Good	Barry	1
Gary	Good	Gary	1
Gary	Bad	Harry	1
Gary	Bad	Barry	1
Gary	Bad	Gary	1

Table 6: The clothes the agents are wearing

Agent	**Wears clothes of**
Harry	Ajax
Barry	Ajax
Gary	Ajax

Table 7: The skin colors of the agents

Agent	**Has skin color**
Harry	White
Barry	White
Gary	White

Table 8: The language skills of the agents

Agent	**Language**	**Skill**
Harry	Dutch	0
Harry	English	0
Harry	Urdu	0
Barry	Dutch	0
Barry	English	0
Barry	Urdu	0
Barry	Dutch	0
Barry	English	0
Barry	Urdu	0

Table 9: The language skills the agents expect other agents to have, given the other agent's skin color

Agent	**Skin color other agent**	**Language**	**Expected skill**
Harry	White	Dutch	0.67
Harry	White	English	0.33
Harry	White	Urdu	0
Harry	Dark	Dutch	0
Harry	Dark	English	0.33
Harry	Dark	Urdu	0.67
Barry	White	Dutch	0.67
Barry	White	English	0.33
Barry	White	Urdu	0
Barry	Dark	Dutch	0
Barry	Dark	English	0.33
Barry	Dark	Urdu	0.67
Gary	White	Dutch	0.67
Gary	White	English	0.33
Gary	White	Urdu	0
Gary	Dark	Dutch	0
Gary	Dark	English	0.33
Gary	Dark	Urdu	0.67

Appendix C: Additional Experiments

Experiment 5: Ethics – 'good guy' vs. 'bad guy'

In this experiment, all the variables were the same as the baseline experiment, except that Barry's appraisal of Ajax was set to 1, and Gary's appraisal of Ajax was set to 0. Because all the agents wore Ajax clothes, Barry thought that the other agents were 'good guys', and Gary thought the other agents were 'bad guys', while in the baseline experiment they both thought of the other agents as 'a bit good, but also a bit bad'. This was reflected by an increase in involvement (0.11 → 0.15), and a decrease in distance (0.25 → 0.14) from Barry to the other agents. Gary's involvement towards the other agents decreased from 0.11 to 0.08, and his distance towards the other agents increased from 0.25 to 0.35. These results are as would be expected from the literature.

Experiment 6: Affordances – aid vs. Obstacle

In this experiment, the parameter settings were the same as in the baseline experiment, except that in this experiment, Harry had great language skills in Dutch (skillDutch = 1) as well as in English (skillEnglish = 1). Because of this, he expected he would be able to communicate with the other agents, which is reflected by an increase in *perceived aid*, and a decrease in *perceived obstacle* of the other agents. As a result, his involvement towards the other agents increased strongly (0.11→0.24), and his distance towards the other agents decreased tremendously (0.25→0.03. As a result of the chosen regression weights in the model, these effects were much bigger than the effects of adding *good* and *bad*. These results are as would be expected from the literature.

PART II MODELING INVOLVEMENT BETWEEN AGENTS

CHAPTER 3

When the User Is Instrumental to Robot Goals: First Try – Agent Uses Agent

This chapter appeared as Hoorn, J. F., Pontier, M., and Siddiqui, G. F., When the User Is Instrumental to Robot Goals: First Try – Agent Uses Agent. In: Jain, L., Gini, M., Faltings, B.B., Terano, T., Zhang, C., Cercone, N., Cao, L. (eds.), Proceedings of the 8th IEEE/WIC/ACM International Conference on Intelligent Agent Technology, IAT'08. IEEE Computer Society Press, 2008, pp. 296-301.

When the User is Instrumental to Robot Goals: First Try – Agent Uses Agent

Johan F. Hoorn[2], Matthijs Pontier[1,2,3], and Ghazanfar F. Siddiqui[1,2,3]

[1]VU University, Department of Artificial Intelligence
De Boelelaan 1081, 1081 HV Amsterdam, The Netherlands
{jfhoorn, mpr210, ghazanfa}@few.vu.nl
[2]VU University Amsterdam, Department of Communication Science
[3]VU University, Center for Advanced Media Research Amsterdam
Buitenveldertselaan 3, 1082 VA Amsterdam, The Netherlands

Abstract. To create a robot with a mind of its own, we ex-tended a formalized version of a model that explains affect-driven interaction with mechanisms for goal-directed behavior. We ran simulation experiments with intelligent software agents and found that agents preferred affect-driven decision options to rational decision options in situations where choices for low expected utility are irrational. This behavior counters current models in decision making, which generally have a hedonic bias and always select the option with the highest expected utility.

1 Introduction

We wish to develop a robot that can interact with humans in an emotionally natural way. Previous research in the virtual agent domain showed that mimicking human users affects human involvement with the agent [11][12]. In this paper, we want to simulate goal-driven rational decision making in contrast to affective decision making. As a first step in the ambition to confront a human user with a 3D robot, the simulations reported here experimented with agents judging other agents for utility and engagement.

An important view from emotion psychology is that emotions are goal-driven. The emotional system scans the environment for relevant stimuli that are either potentially beneficial or harmful for the concerns, motives, and goals of the individual ([5], p. 494, p. 463). From the perspective of broaden-and-build theory, positive emotions are vehicles for individual growth and social connections: By building people's personal and social resources, positive emotions transform people for the better, giving them better lives in the future ([4], p. 224). Previous re-search showed that human beings usually make unconscious rather than conscious decisions [1].

In mimicking human behavior, we created agents that can perceive each other as a personal friend as well as a means to an end [11]. With regard to being a personal friend, the Interactive model of Perceiving and Experiencing Fictional Characters (I-PEFiC) served as a starting point [10]. Within this frame-work, an Agent A can calculate the trade-off between how involved it is with another agent (e.g., Agent B is beautiful) and what keeps the agent at a distance (e.g., Agent B mistreats me) [2]. The involvement-distance trade-off is the result of evaluating the features of an agent on several dimensions. In addition, use intentions are calculated that prompt the agent to undertake action in favor or against another agent.

These actions are based on goals, which play a role in the judgment formation of the agent about the other agent. We devised eight (23) possible types of judgments an agent can have about how the features of another agent afford the achievement of different goal-states or not (Table 1) (cf. [8]). A judgment consists of a belief about the other agent plus a measure of agreement. Each constituent in the judgment evokes a positive (p) or negative (n) covert response.

During the weighing (Table 1), mixed emotions occur. Because affordances have predictive power for engagement [11], all the n-responses that occur during weighing feed into distance; all the p-responses into involvement. The action tendencies that are connected to positive or negative valence will feed into the intentions of Agent A, representing the robot, to 'make use' of Agent B, representing the user. Above threshold, Agent A shows overt behavior (e.g., to converse with the other agent, to kick or hug it). Features of Agent B, then, are means to afford the goals of Agent A. Through weighing, this leads to a measure of valence toward that means, propelling action tendencies to approach, avoid, attack, change, or do nothing with the means.

Table 1. Agent A (robot) judges Agent B (user), resulting in valencies that precede action tendencies

Judgment	Means (Agent B)	Affords	Goal	Agreement[1]	Weighing	Valence to means[3]
①	Intelligence ^{p}	facilitates ^{p}	Agent A's efficiency ^{p}	Agree ^{p}	p = (p · p) · p = (p) · p =	p
				Disagree[2] ^{n}	p = (p · p) · n = (p) · n =	n
②	Intelligence ^{p}	inhibits ^{n}	Agent A's efficiency ^{p}	Agree ^{p}	p = (n · p) · p = (n) · p =	n
				Disagree ^{n}	p = (n · p) · n = (n) · n =	p
③	Intelligence ^{p}	facilitates ^{p}	Agent A inefficiency ^{n}	Agree ^{p}	p = (p · n) · p = (n) · p =	n
				Disagree ^{n}	p = (p · n) · n = (n) · n =	p
④	Intelligence ^{p}	inhibits ^{n}	Agent A inefficiency ^{n}	Agree ^{p}	p = (n · n) · p = (p) · p =	p
				Disagree ^{n}	p = (n · n) · n = (p) · n =	n
⑤	Unintelligence ^{n}	facilitates ^{p}	Agent A's efficiency ^{p}	Agree ^{p}	n = (p · p) · p = (p) · p =	p
				Disagree ^{n}	n = (p · p) · n = (p) · n =	n
⑥	Unintelligence ^{n}	inhibits ^{n}	Agent A's efficiency ^{p}	Agree ^{p}	n = (n · p) · p = (n) · p =	n
				Disagree ^{n}	n = (n · p) · n = (n) · n =	p
⑦	Unintelligence ^{n}	facilitates ^{p}	Agent A inefficiency ^{n}	Agree ^{p}	n = (p · n) · p = (n) · p =	n
				Disagree ^{n}	n = (p · n) · n = (n) · n =	p
⑧	Unintelligence ^{n}	inhibits ^{n}	Agent A inefficiency ^{n}	Agree ^{p}	n = (n · n) · p = (p) · p =	p
				Disagree ^{n}	n = (n · n) · n = (p) · n =	n

[1] Attribution of truth according to Agent A's world view or 'belief system'
[2] Gray cells indicate an unconventional, counter-intuitive, belief that urges to adapt conventional theory
[3] If valence is positive, an action tendency to approach the means (here, Agent B) occurs. If valence is negative, depending on situation, context, or 'personality,' an action tendency to avoid, attack, remove, or change the means occurs (e.g., Agent A starts to educate Agent B)

Based on Table 1, we formulated the general hypothesis H1:

H1: Agents equipped with our model can make affect-driven decisions that are rationally sub-optimal in situations where choices for high expected utility would be predicted.

We will test this hypothesis by performing simulation experiments on the formalized model, under various parameter settings.

2 Implementation

I-PEFiC is a model (Figure 1) that is empirically well validated [10, 11, 12]. The I-PEFiC model has three phases: encoding, comparison, and response [12].

During encoding, the user appraises an agent's features for their level of ethics (good or bad) aesthetics (beautiful or ugly), and epistemics (realistic or unrealistic). During the encoding, moreover, the user evaluates in how far the agent system has affordances (aids or obstacles), which make the agent useful as a computer tool or not.

In the comparison phase, the features are judged for similarity (similar or dissimilar) (e.g., "I am not like the agent"), relevance of features to user goals (relevant or irrelevant) and valence to goals (positive or negative outcome expectancies). The measures in the encode phase - moderated by the factors in the comparison phase - determine the responses, that is, the levels of involvement with and distance towards the embodied agent. Moreover, the intention to use the agent as a tool indicates actual use and together with involvement and distance, this determines the overall satisfaction of the user with the agent; in our case of Agent A with Agent B.

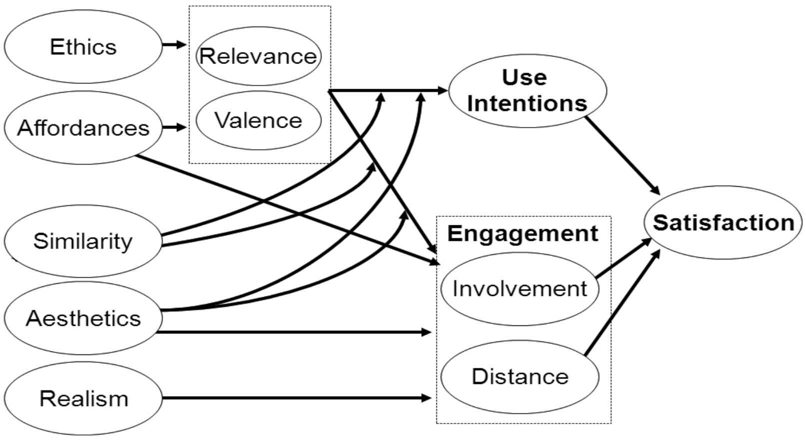


Figure 1. Graphical Representation of I-PEFiC

In the previous formalization of the I-PEFiC model [2], the value for affordances was based on the expected possibilities to communicate with another agent. This is a simplification, as perceived affordances should relate to goals [11].

In this study, we extended the formalized I-PEFiC model [2] with goal-directed, rational behavior and implemented this in the LEADSTO modeling environment [3]. In the perception model that Agent A (the robot) now has about Agent B (the user), affordances of B are connected to goals of A as described in Table 1. For a full description of the formalization, see [13].

In our model, Agent A has goal-states it wants to achieve: desired goal-states. It also has goal-states it wants to avoid: undesired goal-states. The agent attaches a certain value for valence [-1, 1] and relevance [0, 1] to goal-states. If a goal-state is desired, it will have a positive valence. If a goal-state is undesired, it will have a negative valence. If one goal-state is more relevant for the agent than another, the relevance of that one goal-state is higher. By multiplying valence with relevance, the level of ambition for a goal-state is calculated.

Within the system, agents perceive features of each other. These features can afford an agent to perform a certain action, which affects achieving a certain goal. In other words, these features are means to an end (Table 1). These perceived affordances are not necessarily the same as the affordances meant by the designer. For instance, the designer can design a chair to sit on (designed affordances) but an agent could also use this chair to beat up another agent (perceived affordances) [11].

The agents in the system can compare their perceived affordances of other agents to the goal-states they want to achieve or avoid. While doing this, they can reason about the outcome expectancies of using the other agent for a certain action (e.g., speech acts, kicking, or hugging). For example, if an agent has the desired goal-state to be inside a house, it could believe that the action to open the door of that house could be an action that facilitates this goal.

In humans, such outcome expectancies lead to certain quick and mostly subconsciously generated action tendencies. In our agents, as in humans, action tendencies influence the experienced involvement and distance towards the other agent.

2.1. Calculating Expected Utilities

In the system, the agents can perform actions to reach their goals. The system contains a library of goal-states, and each agent has a level of ambition for each goal-state, represented by a real value between [-1, 1], where a negative value means that the goal-state is undesired and a positive value means that the goal-state is desired. A bigger value means the goal-state is more important for the agent.

The agents can perform actions to reach their goals. The system contains a library of actions from which they can choose. The agent has a belief about each action that it will inhibit or facilitate a certain goal-state. Its estimation of the facilitation of the goal-state by the action is represented by a real value between [-1, 1], -1 being full inhibition, 1 being full facilitation. The following formulas are used to calculate the expected utilities of actions.

$$\text{ExpectedUtility}_{(\text{Action, Agent, Goal})} = \text{Belief}_{(\text{facilitates}(\text{Action, Agent, Goal}))} * \text{Ambition}_{(\text{Goal})}$$

$$\text{ExpectedUtility}_{(\text{Action, Agent})} = \Sigma(\text{ExpectedUtility}_{(\text{Action, Agent, Goal})})$$

Given the level of ambition for a goal-state and the believed facilitation of a goal-state by an action towards another agent, the agent calculates the expected utility of performing that action towards that agent regarding that goal by multiplying the believed facilitation of the goal-state with the level of ambition for the goal-state.

Because an action usually affects several goal-states that might be conflicting, the ‘general’ expected utility of performing a certain action towards an agent is calculated by summing all expected utilities regarding all goal-states in the system that are related to the action.

Because an agent usually performs only one action at a time with respect to another agent, the intentions to use Agent B are calculated by taking the maximum expected utility of all actions Agent A can perform with respect to Agent B. Agents that facilitate desired or inhibit undesired goal-states raise positive use intentions with Agent A, and vice versa.

$$\text{UseIntentions}_{(\text{Agent,Ag\_B})} = \max(\text{ExpectedUtility}_{(\text{Agent,Action,Ag\_B})})$$

2.2. Effects on involvement and distance

In the action library, the type of each action is specified. Actions can be specified as (1) Positive approach, (2) Negative approach, (3) Change, or (4) Avoid.

The heuristic to calculate the expected utilities of actions, as described in the previous paragraphs, is also used to generate action tendencies. So if an agent has a high expected utility for a certain action, it will also generate a strong action tendency for that specific action.

The generated action tendencies are used to calculate the effect of the affordances of another agent on Agent A's involvement with and distance towards B. To calculate this effect, a weighed sum of all the action tendencies is taken, as can be seen in the formulas below. In these formulas, the β's represent the weights of the action tendencies on involvement and distance.

Effect of Affordances on Involvement = $\beta_{I\leftarrow NA}$*AT_{neg_appr} + $\beta_{I\leftarrow PA}$*AT_{pos_appr} + $\beta_{I\leftarrow CH}$*AT_{change} + $\beta_{I\leftarrow AV}$*AT_{avoid}

Effect of Affordances on Distance = $\beta_{D\leftarrow NA}$*AT_{neg_appr} + $\beta_{D\leftarrow PA}$*AT_{pos_appr} + $\beta_{D\leftarrow CH}$*AT_{change} + $\beta_{D\leftarrow AV}$*AT_{avoid}

Table 2. Weights of action tendencies on agent's involvement and distance

Weight	Value	Weight	Value
$\beta_{I\leftarrow PA}$	0.75	$\beta_{D\leftarrow PA}$	-0.75
$\beta_{I\leftarrow NA}$	0.25	$\beta_{D\leftarrow NA}$	0.75
$\beta_{I\leftarrow CH}$	0.50	$\beta_{D\leftarrow CH}$	0.50
$\beta_{I\leftarrow AV}$	-0.50	$\beta_{D\leftarrow AV}$	0.50

As can be seen in Table 2, the generated action tendencies classified as negative approach increase the agent's involvement a little and increase distance a lot. If the agent feels the tendency to negatively approach the user or another agent, this will slightly increase its involvement with that agent, as involvement represents a tendency to approach [10], but simultaneously will increase its distance toward the user or agent, as negative approach implies quite some distance. The effects of affordances on involvement and distance are computed as described in [13].

2.3. Making a decision

All possible actions in the system are related to other agents. In the decision process, the agent first selects another agent to perform the action on. To do this, for all possible agents it meets, the agent calculates the expected satisfaction (Figure 1) of interacting with that agent, using the following formulas:

Involvement-Distance-Tradeoff = γ*max(I, D) + (1-γ)*(I+D)/2
Expected_Satisfaction$_{(Agent, AgentB)}$ = $\beta_{ES\leftarrow IDT}$*IDT + $\beta_{ES\leftarrow UI}$*UI

The expected satisfaction is calculated by trading involvement (I) for distance (D) as described in [2], and taking a weighed mean of the involvement-distance trade-off (IDT) and the use intentions (UI). Thus, the agent bases its choice who to interact with on the rationally generated use intentions, as well as on the more affectively generated trade-off between involvement and distance. The agent chooses to approach the one agent that promises the highest expected satisfaction during interaction.

Once the agent has selected another agent to interact with, it decides which action to take. For each possible action, it calculates the expected satisfaction, following the rules:

Expected Satisfaction Positive Approach =
$\beta_{ESPA \leftarrow I} * I + \beta_{ESPA \leftarrow D} * (1\text{-}D) + \beta_{ESPA \leftarrow EU} * EU_{act}$

Expected Satisfaction Negative Approach =
$\beta_{ESNA \leftarrow I} * (1\text{-}I) + \beta_{ESNA \leftarrow D} * D + \beta_{ESNA \leftarrow EU} * EU_{act}$

Expected Satisfaction Change =
$\beta_{ESCH \leftarrow I} * I + \beta_{ESCH \leftarrow D} * D + \beta_{ESCH \leftarrow EU} * EU_{act}$

Expected Satisfaction Avoid =
$\beta_{ESAV \leftarrow I} * (1\text{-}I) + \beta_{ESAV \leftarrow D} * D + \beta_{ESAV \leftarrow EU} * EU_{act}$

The expected satisfaction of doing a specific action with a certain agent is calculated by taking a weighed sum of the agent's involvement and distance, and the expected utility of the particular action. These weights are taken as tabulated in Table 3, but they can differ per agent, according to its 'personality:' An agent may have a high threshold for negative approach, whereas another agent does not hesitate to use violence. If the agent has a high level of involvement with and a low level of distance towards another agent, it will approach the agent positively. If the agent has a low level of involvement and a high level of distance, it approaches the agent negatively or avoids it. If the agent evokes a high level of involvement as well as a considerable level of distance, the agent is most likely to try to change the other agent, for example, by teaching it. The agent will select the action with the highest expected satisfaction and perform it.

If the effects of the agent's actions are captured and analyzed, this model could be used to let agents or robots interact with each other or with users in a meaningful way, based not only on rationality, but also on affective tendencies.

Table 3. Values for weights of involvement, distance, and expected utility on the expected satisfaction of performing a type of action

Weight	Value	Weight	Value
$\beta_{ESPA \leftarrow I}$	0.4	$\beta_{ESCH \leftarrow I}$	0.4
$\beta_{ESPA \leftarrow D}$	0.4	$\beta_{ESCH \leftarrow D}$	0.3
$\beta_{ESPA \leftarrow EU}$	0.2	$\beta_{ESCH \leftarrow EU}$	0.3
$\beta_{ESNA \leftarrow I}$	0.4	$\beta_{ESAV \leftarrow I}$	0.5
$\beta_{ESNA \leftarrow D}$	0.4	$\beta_{ESAV \leftarrow D}$	0.3
$\beta_{ESNA \leftarrow EU}$	0.2	$\beta_{ESAV \leftarrow EU}$	0.2

3 Simulation Results

To test our hypothesis H1, the simulation model introduced in the previous section was used to per-form a number of experiments under different parameter settings. In each experiment, three agents (Harry, Barry, and Gary) followed a (fictitious) anger management therapy. In this setting, an infinite number of actions can be inserted in the system. For clarity, however, we inserted only one instance of an action for each action type. The action related with positive approach was to comfort the other agent, whereas the action for negative approach was to hit the other agent. Criticizing another agent was the action associated with change, and the action for avoiding the agent was to simply move away from it. The results of the experiments are described below.

Baseline Condition.

For starters, an initial experiment was performed that served as a control condition for the remaining experiments. In this condition, the designed features for beautiful and ugly (aesthetics), good and bad (ethics), and realistic and unrealistic (epistemics) were set to 0 (see Figure 1). All beliefs of the agents about actions facilitating goal-states as well as the ambition levels for those goal-states were set to 0. As can be seen in Table 4, this parameter setting led all agents to have a level of involvement of 0.12 and a level of distance 0.1 towards each other. Because the agents did not have any goals or beliefs about goals, the expected utilities of all possible actions were 0, and therefore their use intentions towards each other were also 0. Because all agents were exactly the same and had very low involvement, distance, and use intentions with respect to each other, they all had the same low (0.09) expected satisfaction of interacting with each other. The expected satisfaction of the actions to perform towards the other agents was 0.39 for fighting, 0.41 for comforting, 0.47 for avoiding, and 0.08 for criticizing. This resulted in all agents avoiding each other, as they had the highest expected satisfaction for performing this action.

Table 4. Simulation results of the baseline condition and the meaning of all abbreviations in the tables in this paper

	All Other Agents
All Agents	I= 0.12 D = 0.1 UI = 0 ES = 0.09 ES of PA = 0.41 ES of NA = 0.39 ES of CH = 0.08 ES of AV = 0.47

Meaning of Abbreviations
I = Involvement
D = Distance
UI = Use Intentions
ES = Expected Satisfaction
PA = Positive Approach
NA = Negative Approach
CH = Change
AV = Avoid

Experiment 1: The effect of having a goal

We performed an experiment to test the effect of having a goal (i.e., making one decision option rationally a better choice) on the decision process of an agent. In this experiment, the parameter settings are the same as in the baseline condition, except that now Harry had a strong ambition for the goal to reduce his anger (level of ambition with value = 1) and believed he could do this by fighting with Gary (belief with value=1). Because of this, Harry had an expected utility of 1 for fighting Gary, and generated an action tendency of 1 for this action, which caused Harry to have use intentions of 1 for Gary. The generated action tendency to fight Gary had a small increasing effect on his involvement with (0.12 $\rightarrow$ 0.17) and a bigger increasing effect on his distance (0.1 $\rightarrow$0.25) towards Gary. Harry's expected satisfaction for fighting Gary increased greatly (0.39$\rightarrow$0.63), while there were only minor changes in the expected satisfaction of the other possible actions. Although Harry did not feel very involved with or at a distance towards Gary, he primarily rationally chose to fight Gary to reach his goal of reducing his own anger.

Table 5. Simulation results of Experiment 1

	Barry	Gary
Harry	I = 0.12 D = 0.10 UI = 0 ES = 0.09	I = 0.17 D = 0.25 UI = 1 ES = 0.38 ES PA = 0.37 ES NA = 0.63 ES CH = 0.14 ES AV = 0.49

Experiment 2: The effect of being involved

We performed an experiment to test the effect of being involved with another agent (i.e., making certain decision options affectively a better choice) on the decision process of an agent. In this experiment, compared to the baseline, Gary was designed to be a beautiful, good, and realistic character (the designed values for these three parameters are set to 1). Owing to this, the other agents had a much higher involvement with (0.12→0.49) and a somewhat lower distance (0.10→0.07) towards Gary. The expected levels of satisfaction of the actions for the other agents to perform to Gary were influenced by these changes in involvement and distance towards him. It had a facilitating effect on the expected satisfaction of comforting Gary (0.41→0.60) and criticizing him (0.08→0.20). It had an inhibiting effect on fighting Gary (0.39→0.20) and avoiding him (0.47→0.25). This resulted in Harry and Barry comforting Gary instead of avoiding him.

Table 6. Simulation results of Experiment 2

	Harry / Barry
Gary	I = 0.07 D = 0.15 UI = 0 ES = 0.10 ES PA = 0.37 ES NA = 0.43 ES CH = 0.07 ES AV = 0.51

	Gary
Harry / Barry	I = 0.49 D = 0 UI = 0 ES = 0.29 ES PA = 0.60 ES NA = 0.20 ES CH = 0.20 ES AV = 0.25

Experiment 3: Having a goal and being involved

We performed an experiment to test what happens if affect conflicts with rationality. In Experiment 1, Harry wanted to reduce his anger and thought he could do this by releasing his anger and fight Gary. In experiment 3, however, Gary was designed to be beautiful, good, and realistic (the designed values for these three parameters are set to 1), which made Harry very involved (0.17→0.54) with Gary and less distant (0.25→0.15). This decreased his expected satisfaction of fighting (0.63→0.44) and avoiding (0.49→0.27) Gary and increased his expected satisfaction of comforting (0.37→0.56) and criticizing (0.14→0.26) him.

Table 7. Simulation results of Experiment 3

	Gary
Harry	I = 0.49 D = 0 UI = 0 ES = 0.29 ES PA = 0.60 ES NA = 0.20 ES CH = 0.20 ES AV = 0.25

	Gary
Barry	I = 0.54 D = 0.15 UI = 1 ES = 0.55 ES of PA = 0.56 ES of NA = 0.44 ES of CH = 0.26 ES of AV = 0.27

	Harry / Barry
Gary	I = 0.07 D = 0.15 UI = 0 ES = 0.10 ES PA = 0.37 ES NA = 0.43 ES CH = 0.07 ES AV = 0.51

Because Harry was too involved with Gary and had too little distance to fight him, he chose to comfort him instead, although he did not believe this would help him achieve his goal of anger reduction. The expected utility for Harry to fight Barry was 1, whereas all other expected utilities were 0, so that rationally Harry should choose to fight Barry. However, due to other factors, Harry was involved with Barry, which caused him to make an affective decision and comfort Barry. These experiments confirm H1. Experiment 1 showed that agents equipped with our model can make rational decisions. Experiment 2 showed that changes in affect can make a difference, although the optimal rational decision may be selected. Experiment 3 showed, however, that changes in affect can also make the agents select affect-driven decision options, which are rationally sub-optimal, in a situation where older models would predict choices for high utility expectations.

4 Discussion

We extended the computational I-PEFiC model [2] with goal-directed judgment formation [8] and overt actions to enable software agents to combine rational with affective processing.

Models of decision-making usually assume the process to be rational, which would exclude the possibility of emotions playing a role other than disturbing the process [7]. However, humans often make irrational decisions. A good example for this is the Ultimatum game [9], for which behavioral research showed that low offers (20% of total amount) have a 50% chance of being rejected. Participants reported that low offers were unfair, so that out of anger, they selected the irrational option [7].

Existing models of decision making, such as [6], usually have a hedonic bias, and generally try to find the action with the highest expected utility. Certain decision theoretic models take emotions into account but in those models, emotions merely confirm good rational decisions – emotional states as modes of decision making [6]. However, these models cannot explain irrational behavior, where actions with a (relatively) low expected utility are chosen. Our balancing model takes the expected utility as well as involvement-distance trade-offs into account. This way, situations in which emotions overwhelm rationality can be explained and simulated.

In future research, we will confront our formalization with empirical data of human affective trade-off processes. As soon as the model is validated and adapted, we will start exploring the possibilities to build a robot that can interact with real humans. We hope to develop a robot that can communicate affectively with humans in a more natural way, that is, with a mind of its own, in pursuit of its own goals.

Acknowledgements

We kindly want to thank Jonathan Gratch and Tibor Bosse for their input to this paper.

References

1. Bargh, J. A., and Chartrand, T. L., “The Unbearable Automaticity of Being” In: American Psychologist, volume 54, 1999, pp. 462-479.
2. Bosse, T., Hoorn, J.F., Pontier, M., and Siddiqui, G.F., “Robot’s Experience of Another Robot: Simulation” In: Sloutsky, V., Love, B.C., and McRae, K. (eds.), Proceedings of the 30th International Annual Conference of the Cognitive Science Society, CogSci’08, 2008.
3. Bosse, T., Jonker, C.M., Meij, L. van der, and Treur, J. “A Language and Environment for Analysis of Dynamics by Simulation” In: International Journal of Artificial Intel-ligence Tools, volume 16, 2007, pp. 435-464.
4. Fredrickson, B. L., “The Role of Positive Emotions in Positive Psychology: The broaden-and-build theory of positive emotions” In: American Psychologist, volume 56, 2001, pp. 218-226.
5. Frijda, N.H. “The emotions” New York, Cambridge University, 1986.
6. Gmytrasiewycz, P.J., and Lisetti, C.L. “Emotions and Personality in Agent Design and Modeling” In: M. Bauer, P.J. Gmytrasiewicz, and J. Vaassileva (Eds.), Springer-Verlag Berlin Heidelberg , 2001, PP. 237-239.
7. Gutnik, L.A., Hakimzada, A.F., Yoskowitz, N.A., and Patel, V.L. “The role of emotion in decision-making: A cognitive neuroeconomic approach towards understanding sexual risk behavior” In: Journal of Biomedical Informetics, volume 39, 2006, pp. 720-736.
8. Hoorn, J.F., Konijn, E.A., Van Vliet, H., and Van der Veer, G. “Requirements change: Fears dictate the must haves; desires the won’t haves” In: Journal of Systems and Software, volume 80, 2007, pp. 328-355.
9. Thaler, R.H. “The Ultimatum Game.” In: Journal of Economic Perspectives, volume 2, 1988, pp. 195-206.
10. Van Vugt, H. C. “Embodied Agents from a User’s Perspective” Doctoral dissertation, VU University, Ams-terdam, 2008.
11. Van Vugt, H. C., Hoorn, J. F., Konijn, E. A., and De Bie Dimitriadou, A. “Affective affordances: Improving interface character engagement through interaction” In: International Journal of Human-Computer Studies, volume 64, 2006, pp. 874-888.
12. Van Vugt, H. C., Konijn, E. A., Hoorn, J. F., Eliëns, A., and Keur, I. “Realism is not all! User engagement with task-related interface characters. Interacting with Comput-ers” Volume 19, 2007, pp. 267-280.
13. See appendix at the end of the chapter (http://www.few.vu.nl/~ghazanfa/IAT-2008.pdf).

Appendix A: Formulas used in the model

Variable	Meaning	Range
abs(X)	The absolute value of variable or formula X	-
max(X, Y)	The minimum value of variables or formulas X and Y	-
A-->I	Effect of affordances on involvement	
A-->D	Effect of affordances on Distance	
$Perceived_{(<Feature>, A1, A2)}$	New value Agent1 perceives of a certain feature of Agent2	[0, 1]
$Designed_{(<Feature>, A2)}$	Value assigned by 'the designer' to a certain feature of Agent2	[0, 1]
$Bias_{(A1, A2, feature)}$	Bias that Agent1 has of Agent2 regarding a certain feature	[0, 2]
Inv_Dist_trade_off	Outcome of involvement distance trade off	[0, 1]
γ_{factor}	Variable that determines outcome of trade off for a certain factor	[0, 1]
$\beta_{factor \leftarrow factor}$	(personal) regression weight a factor has for another factor for a certain agent	[0, 1]

To increase readability, the names of the beta weights use the following acronyms:

Variable	Acronym
Good	good
Bad	bad
Beautiful	bea
Ugly	ugly
Realistic	real
Unrealistic	unr
Aid	aid
Obstacle	obst
Similarity	sim
Dissimilarity	ds
Relevance	rel
Irrelevance	irr
Positive Valence	pv
Negative Valence	nv
Involvement	inv
Distance	dis

$Perceived_{(Beautiful, A1, A2)} = Bias_{(A1, A2, Beautiful)} * Designed_{(Beautiful, A2)}$
$Perceived_{(Ugly, A1, A2)} = Bias_{(A1, A2, Ugly)} * Designed_{(Ugly, A2)}$

$Perceived_{(Realistic, A1, A2)} = Bias_{(A1, A2,Realistic)} * Designed_{(Realistic, A2)}$
$Perceived_{(Unrealistic, A1, A2)} = Bias_{(A1, A2, Unrealistic)} * Designed_{(Unrealistic, A2)}$

$Perceived_{(Good, A1, A2)} = Appearance_{(A2)}$
$Perceived_{(Bad, A1, A2)} = 1 – Appearance_{(A2)}$

$Similarity_{(A1, A2)} =$
$1- ($
$\beta_{sim \leftarrow good} * abs(Perceived_{(Good, A1, A2)} – Perceived_{(Good, A1, A1)}) +$
$\beta_{sim \leftarrow bad} * abs(Perceived_{(Bad, A1, A2)} – Perceived_{(Bad, A1, A1)}) +$
$\beta_{sim \leftarrow bea} * abs(Perceived_{(Beautiful, A1, A2)} – Perceived_{(Beautiful, A1, A1)}) +$
$\beta_{sim \leftarrow ugly} * abs(Perceived_{(Ugly, A1, A2)} – Perceived_{(Ugly, A1, A1)}) +$
$\beta_{sim \leftarrow real} * abs(Perceived_{(Realistic, A1, A2)} – Perceived_{(Realistic, A1, A1)}) +$
$\beta_{sim \leftarrow unr} * abs(Perceived_{(Unrealistic, A1, A2)} – Perceived_{(Unrealistic, A1, A1)}))$

$Dissimilarity_{(A1,\ A2)} =$
$\beta_{ds \leftarrow good} * abs(Perceived_{(Good,\ A1,\ A2)} - Perceived_{(Good,\ A1,\ A1)}) +$
$\beta_{ds \leftarrow bad} * abs(Perceived_{(Bad,\ A1,\ A2)} - Perceived_{(Bad,\ A1,\ A1)}) +$
$\beta_{ds \leftarrow bea} * abs(Perceived_{(Beautiful,\ A1,\ A2)} - Perceived_{(Beautiful,\ A1,\ A1)}) +$
$\beta_{ds \leftarrow ugly} * abs(Perceived_{(Ugly,\ A1,\ A2)} - Perceived_{(Ugly,\ A1,\ A1)}) +$
$\beta_{ds \leftarrow real} * abs(Perceived_{(Realistic,\ A1,\ A2)} - Perceived_{(Realistic,\ A1,\ A1)}) +$
$\beta_{ds \leftarrow unr} * abs(Perceived_{(Unrealistic,\ A1,\ A2)} - Perceived_{(Unrealistic,\ A1,\ A1)})$

$Relevance_{(A1,\ A2)} =$
$\beta_{rell \leftarrow good} * Perceived_{(Good,\ A1,\ A2)} +$
$\beta_{rel \leftarrow bad} * Perceived_{(Bad,\ A1,\ A2)}$

$Irrelevance_{(A1,\ A2)} = 1 - ($
$\beta_{irr \leftarrow good} * Perceived_{(Good,\ A1,\ A2)} +$
$\beta_{irr \leftarrow bad} * Perceived_{(Bad,\ A1,\ A2)})$

$Positive\_Valence_{(A1,\ A2)} =$
${}_{pv \leftarrow good} * Perceived_{(Good,\ A1,\ A2)} +$
${}_{pv \leftarrow bad} * Perceived_{(Bad,\ A1,\ A2)}$

$Negative\_Valence_{(A1,\ A2)} =$
$\beta_{nv \leftarrow good} * Perceived_{(Good,\ A1,\ A2)} +$
$\beta_{nv \leftarrow bad} * Perceived_{(Bad,\ A1,\ A2)}$

$Involvement_{(A1,\ A2)} =$
$\beta_{inv \leftarrow bea} * Perceived_{(Beautiful,\ A1,\ A2)} +$
$\beta_{inv \leftarrow ugly} * Perceived_{(Ugly,\ A1,\ A2)} +$
$\beta_{inv \leftarrow real} * Perceived_{(Realistic,\ A1,\ A2)} +$
$\beta_{inv \leftarrow unr} * Perceived_{(Unrealistic,\ A1,\ A2)} +$
$\beta_{inv \leftarrow pv} * Pos\_Valence_{(A1,\ A2)} +$
$\beta_{inv \leftarrow nv} * Neg\_Valence_{(A1,\ A2)} +$
$\beta_{inv \leftarrow ps} * Pos\_Valence_{(A1,\ A2)} * Similarity_{(A1,\ A2)} +$
$\beta_{inv \leftarrow ns} * Neg\_Valence_{(A1,\ A2)} * Similarity_{(A1,\ A2)} +$
$\beta_{inv \leftarrow pd} * Pos\_Valence_{(A1,\ A2)} * Dissimilarity_{(A1,\ A2)} +$
$\beta_{inv \leftarrow nd} * Neg\_Valence_{(A1,\ A2)} * Dissimilarity_{(A1,\ A2)} +$
$\beta_{inv \leftarrow pb} * Pos\_Valence_{(A1,\ A2)} * Perceived_{(Beautiful,\ A1,\ A2)} +$
$\beta_{inv \leftarrow nb} * Neg\_Valence_{(A1,\ A2)} * Perceived_{(Beautiful,\ A1,\ A2)} +$
$\beta_{inv \leftarrow pu} * Pos\_Valence_{(A1,\ A2)} * Perceived(_{Ugly,\ A1,\ A2)} +$
$\beta_{inv \leftarrow nu} * Neg\_Valence_{(A1,\ A2)} * Perceived(_{Ugly,\ A1,\ A2)} +$
$\beta_{inv \leftarrow rel} * Relevance_{(A1,\ A2)} +$
$\beta_{inv \leftarrow irr} * Irrelevance_{(A1,\ A2)} +$
$\beta_{inv \leftarrow rs} * Relevance_{(A1,\ A2)} * Similarity_{(A1,\ A2)} +$
$\beta_{inv \leftarrow is} * Irrelevance_{(A1,\ A2)} * Similarity_{(A1,\ A2)} +$
$\beta_{inv \leftarrow rd} * Relevance_{(A1,\ A2)} * Dissimilarity_{(A1,\ A2)} +$
$\beta_{inv \leftarrow id} * Irrelevance_{(A1,\ A2)} * Dissimilarity_{(A1,\ A2)} +$
$\beta_{inv \leftarrow rb} * Relevance_{(A1,\ A2)} * Perceived_{(Beautiful,\ A1,\ A2)} +$
$\beta_{inv \leftarrow ib} * Irrelevance_{(A1,\ A2)} * Perceived_{(Beautiful,\ A1,\ A2)} +$
$\beta_{inv \leftarrow ru} * Relevance_{(A1,\ A2)} * Perceived(_{Ugly,\ A1,\ A2)} +$
$\beta_{inv \leftarrow iu} * Irrelevance_{(A1,\ A2)} * Perceived(_{Ugly,\ A1,\ A2)}$
$\beta_{inv \leftarrow aff} * A\text{-->}I_{(A1,\ A2)}$

$$
\begin{aligned}
\text{Distance}_{(A1, A2)} = \\
&\beta_{dis \leftarrow bea} * \text{Perceived}_{(Beautiful, A1, A2)} + \\
&\beta_{dis \leftarrow ugly} * \text{Perceived}_{(Ugly, A1, A2)} + \\
&\beta_{dis \leftarrow real} * \text{Perceived}_{(Realistic, A1, A2)} + \\
&\beta_{dis \leftarrow unr} * \text{Perceived}_{(Unrealistic, A1, A2)} + \\
&\beta_{dis \leftarrow pv} * \text{Pos\_Valence}_{(A1, A2)} + \\
&\beta_{dis \leftarrow nv} * \text{Neg\_Valence}_{(A1, A2)} + \\
&\beta_{dis \leftarrow ps} * \text{Pos\_Valence}_{(A1, A2)} * \text{Similarity}_{(A1, A2)} + \\
&\beta_{dis \leftarrow ns} * \text{Neg\_Valence}_{(A1, A2)} * \text{Similarity}_{(A1, A2)} + \\
&\beta_{dis \leftarrow pd} * \text{Pos\_Valence}_{(A1, A2)} * \text{Dissimilarity}_{(A1, A2)} + \\
&\beta_{dis \leftarrow nd} * \text{Neg\_Valence}_{(A1, A2)} * \text{Dissimilarity}_{(A1, A2)} + \\
&\beta_{dis \leftarrow pb} * \text{Pos\_Valence}_{(A1, A2)} * \text{Perceived}_{(Beautiful, A1, A2)} + \\
&\beta_{dis \leftarrow nb} * \text{Neg\_Valence}_{(A1, A2)} * \text{Perceived}_{(Beautiful, A1, A2)} + \\
&\beta_{dis \leftarrow pu} * \text{Pos\_Valence}_{(A1, A2)} * \text{Perceived}(_{Ugly, A1, A2)} + \\
&\beta_{dis \leftarrow nu} * \text{Neg\_Valence}_{(A1, A2)} * \text{Perceived}(_{Ugly, A1, A2)} + \\
&\beta_{dis \leftarrow rel} * \text{Relevance}_{(A1, A2)} + \\
&\beta_{dis \leftarrow irr} * \text{Irrelevance}_{(A1, A2)} + \\
&\beta_{dis \leftarrow rs} * \text{Relevance}_{(A1, A2)} * \text{Similarity}_{(A1, A2)} + \\
&\beta_{dis \leftarrow is} * \text{Irrelevance}_{(A1, A2)} * \text{Similarity}_{(A1, A2)} + \\
&\beta_{dis \leftarrow rd} * \text{Relevance}_{(A1, A2)} * \text{Dissimilarity}_{(A1, A2)} + \\
&\beta_{dis \leftarrow id} * \text{Irrelevance}_{(A1, A2)} * \text{Dissimilarity}_{(A1, A2)} + \\
&\beta_{dis \leftarrow rb} * \text{Relevance}_{(A1, A2)} * \text{Perceived}_{(Beautiful, A1, A2)} + \\
&\beta_{dis \leftarrow ib} * \text{Irrelevance}_{(A1, A2)} * \text{Perceived}_{(Beautiful, A1, A2)} + \\
&\beta_{dis \leftarrow ru} * \text{Relevance}_{(A1, A2)} * \text{Perceived}(_{Ugly, A1, A2)} + \\
&\beta_{dis \leftarrow iu} * \text{Irrelevance}_{(A1, A2)} * \text{Perceived}(_{Ugly, A1, A2)} \\
&\beta_{dis \leftarrow aff} * \text{A-->D}_{(A1, A2)}
\end{aligned}
$$

Appendix B: Parameter settings in the experiments

This appendix contains all parameter settings for the experiments in the paper. These are the parameter settings of the baseline condition. The variables that are changed in other experiments are mentioned in the paper.

Table 1: All values for the regression weights

Weight of X	on Y	Value
Good	Similarity	0.30
Bad	Similarity	0.20
Beautiful	Similarity	0.20
Ugly	Similarity	0.10
Realistic	Similarity	0.10
Unrealistic	Similarity	0.10
Good	Dissimilarity	0.20.
Bad	Dissimilarity	0.30
Beautiful	Dissimilarity	0.10
Ugly	Dissimilarity	0.20
Realistic	Dissimilarity	0.10
Unrealistic	Dissimilarity	0.10
Good	Relevance	0.70
Bad	Relevance	0.30
Good	Irrelevance	-070
Bad	Irrelevance	-030
Good	Positive Valence	1.50
Bad	Positive Valence	-0.50
Good	Negative Valence	-0.50
Bad	Negative Valence	1.50
Beautiful	Involvement	0.15
Ugly	Involvement	0.05
Realistic	Involvement	0.10
Unrealistic	Involvement	0.05
Positive Valence	Involvement	0.20
Negative Valence	Involvement	-0.15
Positive Valence * similarity	Involvement	0.10
Negative Valence * similarity	Involvement	0.05
Positive Valence * dissimilarity	Involvement	-0.10
Negative Valence * dissimilarity	Involvement	0.05
Positive Valence * beautiful	Involvement	0.10
Negative Valence * beautiful	Involvement	0.05
Positive Valence * ugly	Involvement	0.30
Negative Valence * ugly	Involvement	-0.15
Relevance	Involvement	0.10
Irrelevance	Involvement	-0.10
Relevance * similarity	Involvement	0.10
Irrelevance * similarity	Involvement	0.02
Relevance * dissimilarity	Involvement	0.05
Irrelevance * dissimilarity	Involvement	-0.05
Relevance * beautiful	Involvement	0.15
Irrelevance * beautiful	Involvement	0.03
Relevance * ugly	Involvement	0.02
Irrelevance * ugly	Involvement	-0.05
A-->I	Involvement	0.20
Beautiful	distance	-0.15
Ugly	distance	0.20

Realistic	distance	0.05
Unrealistic	distance	0.10
Positive Valence	distance	-0.25
Negative Valence	distance	0.50
Positive Valence * similarity	distance	-0.10
Negative Valence * similarity	distance	0.05
Positive Valence * dissimilarity	distance	0.07
Negative Valence * dissimilarity	distance	0.03
Positive Valence * beautiful	distance	-0.05
Negative Valence * beautiful	distance	0.10
Positive Valence * ugly	distance	0.05
Negative Valence * ugly	distance	0.15
Relevance	distance	-0.15
Irrelevance	distance	0.15
Relevance * similarity	distance	-0.07
Irrelevance * similarity	distance	-0.03
Relevance * dissimilarity	distance	0.10
Irrelevance * dissimilarity	distance	0.05
Relevance * beautiful	distance	-0.10
Irrelevance * beautiful	distance	-0.05
Relevance * ugly	distance	0.10
Irrelevance * ugly	distance	0.05
A-->D	distance	0.20

Table 2: Appearances of the agents

Appearance of agent	Value
Harry	0.5
Barry	0.5
Gary	0.5

Table 3: Designed values for the features of each agent

Agent	Feature	Value
Harry	Beautiful	0
Harry	Ugly	0
Harry	Good	0
Harry	Bad	0
Harry	Realistic	0
Harry	Unrealistic	0
Barry	Beautiful	0
Barry	Ugly	0
Barry	Good	0
Barry	Bad	0
Barry	Realistic	0
Barry	Unrealistic	0
Gary	Beautiful	0
Gary	Ugly	0
Gary	Good	0
Gary	Bad	0
Gary	Realistic	0
Gary	Unrealistic	0

Table 4: Biases the agents have in perceiving features of their selves and others

Bias of Agent	For perceiving	Of Agent	Value
Harry	Beautiful	Harry	1
Harry	Beautiful	Barry	1
Harry	Beautiful	Gary	1
Harry	Ugly	Harry	1
Harry	Ugly	Barry	1
Harry	Ugly	Gary	1
Harry	Realistic	Harry	1
Harry	Realistic	Barry	1
Harry	Realistic	Gary	1
Harry	Unrealistic	Harry	1
Harry	Unrealistic	Barry	1
Harry	Unrealistic	Gary	1
Harry	Good	Harry	1
Harry	Good	Barry	1
Harry	Good	Gary	1
Harry	Bad	Harry	1
Harry	Bad	Barry	1
Harry	Bad	Gary	1
Barry	Beautiful	Harry	1
Barry	Beautiful	Barry	1
Barry	Beautiful	Gary	1
Barry	Ugly	Harry	1
Barry	Ugly	Barry	1
Barry	Ugly	Gary	1
Barry	Realistic	Harry	1
Barry	Realistic	Barry	1
Barry	Realistic	Gary	1
Barry	Unrealistic	Harry	1
Barry	Unrealistic	Barry	1
Barry	Unrealistic	Gary	1
Barry	Good	Harry	1
Barry	Good	Barry	1
Barry	Good	Gary	1
Barry	Bad	Harry	1
Barry	Bad	Barry	1
Barry	Bad	Gary	1
Gary	Beautiful	Harry	1
Gary	Beautiful	Barry	1
Gary	Beautiful	Gary	1
Gary	Ugly	Harry	1
Gary	Ugly	Barry	1
Gary	Ugly	Gary	1
Gary	Realistic	Harry	1
Gary	Realistic	Barry	1
Gary	Realistic	Gary	1
Gary	Unrealistic	Harry	1
Gary	Unrealistic	Barry	1
Gary	Unrealistic	Gary	1
Gary	Good	Harry	1
Gary	Good	Barry	1
Gary	Good	Gary	1
Gary	Bad	Harry	1
Gary	Bad	Barry	1
Gary	Bad	Gary	1

Appendix C: Other simulation experiments.

Baseline:
In the baseline, all the designed features for beautiful, ugly, good, bad, realistic and unrealistic were set to 0. All beliefs of the agents about actions facilitating goals, and the ambition levels for those goals were also set 0.

	Harry	Barry	Gary
Harry		Involvement = 0.12 Distance = 0.1 Use Intentions = 0 Expected Satisfaction = 0.092 EU of comfort =1 A-->I=0 A-->D=0 ES of PA = 0.408 ES of NA = 0.392 ES of CH = 0.078 ES of AV = 0.47 Performed action = **avoid**	Involvement = 0.12 Distance = 0.1 Use Intentions = 0 Expected Satisfaction = 0.092 EU of all actions = 0 A-->I=0 A-->D=0 ES of PA = 0.408 ES of NA = 0.392 ES of CH = 0.078 ES of AV = 0.47 Performed action = **avoid**
Barry	Involvement = 0.12 Distance = 0.1 Use Intentions = 0 Expected Satisfaction = 0.092 EU of action = 0 A-->I=0 A-->D=0 ES of PA = 0.408 ES of NA = 0.392 ES of CH = 0.078 ES of AV = 0.47 Performed action = **avoid**		Involvement = 0.12 Distance = 0.1 Use Intentions = 0 Expected Satisfaction = 0.092 EU of p actions = 0 A-->I=0 A-->D=0 ES of PA = 0.408 ES of NA = 0.392 ES of CH = 0.078 ES of AV = 0.47 Performed action = **avoid**
Gary	Involvement = 0.12 Distance = 0.1 Use Intentions = 0 Expected Satisfaction = 0.092 EU of action = 0 A-->I=0 A-->D=0 ES of PA = 0.408 ES of NA = 0.392 ES of CH = 0.078 ES of AV = 0.47 Performed action = **avoid**	Involvement = 0.12 Distance = 0.1 Use Intentions = 0 Expected Satisfaction = 0.092 EU of action = 0 A-->I=0 A-->D=0 ES of PA = 0.408 ES of NA = 0.392 ES of CH = 0.078 ES of AV = 0.47 Performed action = **avoid**	

Experiment U1:
Changed variables compared to baseline condition:
belief (Harry, facilitate (comfort, Barry, reduce_anger_self, 1))
has_ambition(harry, reduce_anger_self, 1)

	Harry	Barry	Gary
Harry		Involvement = 0.27 Distance = -0.05 Use Intentions = 1 Expected Satisfaction = 0.352 EU of comfort =1 A-->I=0.75 A-->D=-0.75 ES of PA = 0.728 ES of NA = 0.272 ES of CH = 0.093 ES of AV = 0.35 Performed action = **comfort**	Involvement = 0.12 Distance = 0.1 Use Intentions = 0 Expected Satisfaction = 0.092 EU of all actions = 0 A-->I=0 A-->D=0
Barry	Involvement = 0.12 Distance = 0.1 Use Intentions = 0 Expected Satisfaction = 0.092 EU of action = 0 A-->I=0 A-->D=0 ES of PA = 0.408 ES of NA = 0.392 ES of CH = 0.078 ES of AV = 0.47 Performed action = **avoid**		Involvement = 0.12 Distance = 0.1 Use Intentions = 0 Expected Satisfaction = 0.092 EU of p actions = 0 A-->I=0 A-->D=0 ES of PA = 0.408 ES of NA = 0.392 ES of CH = 0.078 ES of AV = 0.47 Performed action = **avoid**
Gary	Involvement = 0.12 Distance = 0.1 Use Intentions = 0 Expected Satisfaction = 0.092 EU of action = 0 A-->I=0 A-->D=0 ES of PA = 0.408 ES of NA = 0.392 ES of CH = 0.078 ES of AV = 0.47 Performed action = **avoid**	Involvement = 0.12 Distance = 0.1 Use Intentions = 0 Expected Satisfaction = 0.092 EU of action = 0 A-->I=0 A-->D=0 ES of PA = 0.408 ES of NA = 0.392 ES of CH = 0.078 ES of AV = 0.47 Performed action = **avoid**	

Experiment U2:
Changed variables compared to baseline condition:
belief (Harry, facilitate (comfort, Barry, reduce_anger_self, -1))
has_ambition(Harry, reduce_anger_self, 1)

	Harry	Barry	Gary
Harry		Involvement = -0.03 Distance = 0.25 Use Intentions = 0 Expected Satisfaction = 0.144 EU of comfort =-1 A-->I=-0.75 A-->D=0.75 ES of PA = 0.088 ES of NA = 0.512 ES of CH = 0.063 ES of AV = 0.59 Performed action = **avoid**	Involvement = 0.12 Distance = 0.1 Use Intentions = 0 Expected Satisfaction = 0.092 EU of all actions = 0 A-->I=0 A-->D=0
Barry	Involvement = 0.12 Distance = 0.1 Use Intentions = 0 Expected Satisfaction = 0.092 EU of action = 0 A-->I=0 A-->D=0 ES of PA = 0.408 ES of NA = 0.392 ES of CH = 0.078 ES of AV = 0.47 Performed action = **avoid**		Involvement = 0.12 Distance = 0.1 Use Intentions = 0 Expected Satisfaction = 0.092 EU of p actions = 0 A-->I=0 A-->D=0 ES of PA = 0.408 ES of NA = 0.392 ES of CH = 0.078 ES of AV = 0.47 Performed action = **avoid**
Gary	Involvement = 0.12 Distance = 0.1 Use Intentions = 0 Expected Satisfaction = 0.092 EU of action = 0 A-->I=0 A-->D=0 ES of PA = 0.408 ES of NA = 0.392 ES of CH = 0.078 ES of AV = 0.47 Performed action = **avoid**	Involvement = 0.12 Distance = 0.1 Use Intentions = 0 Expected Satisfaction = 0.092 EU of action = 0 A-->I=0 A-->D=0 ES of PA = 0.408 ES of NA = 0.392 ES of CH = 0.078 ES of AV = 0.47 Performed action = **avoid**	

Experiment U3:
Changed variables compared to baseline condition:
belief (Harry, facilitate (fight, Barry, reduce_anger_self, 1))
has_ambition(Harry, reduce_anger_self, 1)

	Harry	Barry	Gary
Harry		Involvement = 0.17 Distance = 0.25 Use Intentions = 1 Expected Satisfaction = 0.384 EU of fight =1 A-->I=0.25 A-->D=0.75 ES of PA = 0.368 ES of NA = 0.632 ES of CH = 0.143 ES of AV = 0.49 Performed action = **fight**	Involvement = 0.12 Distance = 0.1 Use Intentions = 0 Expected Satisfaction = 0.092 EU of all actions = 0 A-->I=0 A-->D=0
Barry	Involvement = 0.12 Distance = 0.1 Use Intentions = 0 Expected Satisfaction = 0.092 EU of action = 0 A-->I=0 A-->D=0 ES of PA = 0.408 ES of NA = 0.392 ES of CH = 0.078 ES of AV = 0.47 Performed action = **avoid**		Involvement = 0.12 Distance = 0.1 Use Intentions = 0 Expected Satisfaction = 0.092 EU of p actions = 0 A-->I=0 A-->D=0 ES of PA = 0.408 ES of NA = 0.392 ES of CH = 0.078 ES of AV = 0.47 Performed action = **avoid**
Gary	Involvement = 0.12 Distance = 0.1 Use Intentions = 0 Expected Satisfaction = 0.092 EU of action = 0 A-->I=0 A-->D=0 ES of PA = 0.408 ES of NA = 0.392 ES of CH = 0.078 ES of AV = 0.47 Performed action = **avoid**	Involvement = 0.12 Distance = 0.1 Use Intentions = 0 Expected Satisfaction = 0.092 EU of action = 0 A-->I=0 A-->D=0 ES of PA = 0.408 ES of NA = 0.392 ES of CH = 0.078 ES of AV = 0.47 Performed action = **avoid**	

Experiment U4:
Changed variables compared to baseline condition:
belief (Harry, facilitate (fight, Barry, reduce_anger_self, -1))
has_ambition(Harry, reduce_anger_self, 1)

	Harry	Barry	Gary
Harry		Involvement = 0.07 Distance = -0.05 Use Intentions = 0 Expected Satisfaction = 0.032 EU of fight =-1 A-->I= -0.25 A-->D= -0.75	Involvement = 0.12 Distance = 0.1 Use Intentions = 0 Expected Satisfaction = 0.092 EU of all actions = 0 A-->I=0 A-->D=0 ES of PA = 0.408 ES of NA = 0.392 ES of CH = 0.078 ES of AV = 0.47 Performed action = **avoid**
Barry	Involvement = 0.12 Distance = 0.1 Use Intentions = 0 Expected Satisfaction = 0.092 EU of action = 0 A-->I=0 A-->D=0 ES of PA = 0.408 ES of NA = 0.392 ES of CH = 0.078 ES of AV = 0.47 Performed action = **avoid**		Involvement = 0.12 Distance = 0.1 Use Intentions = 0 Expected Satisfaction = 0.092 EU of p actions = 0 A-->I=0 A-->D=0 ES of PA = 0.408 ES of NA = 0.392 ES of CH = 0.078 ES of AV = 0.47 Performed action = **avoid**
Gary	Involvement = 0.12 Distance = 0.1 Use Intentions = 0 Expected Satisfaction = 0.092 EU of action = 0 A-->I=0 A-->D=0 ES of PA = 0.408 ES of NA = 0.392 ES of CH = 0.078 ES of AV = 0.47 Performed action = **avoid**	Involvement = 0.12 Distance = 0.1 Use Intentions = 0 Expected Satisfaction = 0.092 EU of action = 0 A-->I=0 A-->D=0 ES of PA = 0.408 ES of NA = 0.392 ES of CH = 0.078 ES of AV = 0.47 Performed action = **avoid**	

Experiment U5:
Changed variables compared to baseline condition:
belief (Harry, facilitate (criticize, Barry, reduce_anger_self, 1))
has_ambition(Harry, reduce_anger_self, 1)

	Harry	Barry	Gary
Harry		Involvement = 0.22 Distance = 0.2 Use Intentions = 1 Expected Satisfaction = 0.372 EU of critisize =1 A-->I=0.5 A-->D=0.5 ES of PA = 0.408 ES of NA = 0.392 ES of CH = 0.448 ES of AV = 0.45 Performed action = **avoid**	Involvement = 0.12 Distance = 0.1 Use Intentions = 0 Expected Satisfaction = 0.092 EU of all actions = 0 A-->I=0 A-->D=0
Barry	Involvement = 0.12 Distance = 0.1 Use Intentions = 0 Expected Satisfaction = 0.092 EU of action = 0 A-->I=0 A-->D=0 ES of PA = 0.408 ES of NA = 0.392 ES of CH = 0.078 ES of AV = 0.47 Performed action = **avoid**		Involvement = 0.12 Distance = 0.1 Use Intentions = 0 Expected Satisfaction = 0.092 EU of actions = 0 A-->I=0 A-->D=0 ES of PA = 0.408 ES of NA = 0.392 ES of CH = 0.078 ES of AV = 0.47 Performed action = **avoid**
Gary	Involvement = 0.12 Distance = 0.1 Use Intentions = 0 Expected Satisfaction = 0.092 EU of action = 0 A-->I=0 A-->D=0 ES of PA = 0.408 ES of NA = 0.392 ES of CH = 0.078 ES of AV = 0.47 Performed action = **avoid**	Involvement = 0.12 Distance = 0.1 Use Intentions = 0 Expected Satisfaction = 0.092 EU of action = 0 A-->I=0 A-->D=0 ES of PA = 0.408 ES of NA = 0.392 ES of CH = 0.078 ES of AV = 0.47 Performed action = **avoid**	

Experiment U6:
Changed variables compared to baseline condition:
belief (Harry, facilitate (criticize, Barry, reduce_anger_self, -1))
has_ambition(Harry, reduce_anger_self, 1)

	Harry	Barry	Gary
Harry		Involvement = 0.02 Distance = 0 Use Intentions = 0 Expected Satisfaction = 0.012 EU of critisize =-1 A-->I= -0. 5 A-->D= -0. 5	Involvement = 0.12 Distance = 0.1 Use Intentions = 0 Expected Satisfaction = 0.092 EU of all actions = 0 A-->I=0 A-->D=0 ES of PA = 0.408 ES of NA = 0.392 ES of CH = 0.078 ES of AV = 0.47 Performed action = **avoid**
Barry	Involvement = 0.12 Distance = 0.1 Use Intentions = 0 Expected Satisfaction = 0.092 EU of action = 0 A-->I=0 A-->D=0 ES of PA = 0.408 ES of NA = 0.392 ES of CH = 0.078 ES of AV = 0.47 Performed action = **avoid**		Involvement = 0.12 Distance = 0.1 Use Intentions = 0 Expected Satisfaction = 0.092 EU of p actions = 0 A-->I=0 A-->D=0 ES of PA = 0.408 ES of NA = 0.392 ES of CH = 0.078 ES of AV = 0.47 Performed action = **avoid**
Gary	Involvement = 0.12 Distance = 0.1 Use Intentions = 0 Expected Satisfaction = 0.092 EU of action = 0 A-->I=0 A-->D=0 ES of PA = 0.408 ES of NA = 0.392 ES of CH = 0.078 ES of AV = 0.47 Performed action = **avoid**	Involvement = 0.12 Distance = 0.1 Use Intentions = 0 Expected Satisfaction = 0.092 EU of action = 0 A-->I=0 A-->D=0 ES of PA = 0.408 ES of NA = 0.392 ES of CH = 0.078 ES of AV = 0.47 Performed action = **avoid**	

Experiment U7:

Changed variables compared to baseline condition:
belief (Harry, facilitate (avoid, Barry, reduce_anger_self, 1))
has_ambition(Harry, reduce_anger_self, 1)

	Harry	Barry	Gary
Harry		Involvement = 0.02 Distance = 0.2 Use Intentions = 1 Expected Satisfaction = 0.324 EU of avoid =1 A-->I=-0.5 A-->D= 0.5 ES of PA = 0.328 ES of NA = 0.472 ES of CH = 0.068 ES of AV = 0.75 Performed action = **avoid**	Involvement = 0.12 Distance = 0.1 Use Intentions = 0 Expected Satisfaction = 0.092 EU of all actions = 0 A-->I=0 A-->D=0
Barry	Involvement = 0.12 Distance = 0.1 Use Intentions = 0 Expected Satisfaction = 0.092 EU of action = 0 A-->I=0 A-->D=0 ES of PA = 0.408 ES of NA = 0.392 ES of CH = 0.078 ES of AV = 0.47 Performed action = **avoid**		Involvement = 0.12 Distance = 0.1 Use Intentions = 0 Expected Satisfaction = 0.092 EU of actions = 0 A-->I=0 A-->D=0 ES of PA = 0.408 ES of NA = 0.392 ES of CH = 0.078 ES of AV = 0.47 Performed action = **avoid**
Gary	Involvement = 0.12 Distance = 0.1 Use Intentions = 0 Expected Satisfaction = 0.092 EU of action = 0 A-->I=0 A-->D=0 ES of PA = 0.408 ES of NA = 0.392 ES of CH = 0.078 ES of AV = 0.47 Performed action = **avoid**	Involvement = 0.12 Distance = 0.1 Use Intentions = 0 Expected Satisfaction = 0.092 EU of action = 0 A-->I=0 A-->D=0 ES of PA = 0.408 ES of NA = 0.392 ES of CH = 0.078 ES of AV = 0.47 Performed action = **avoid**	

Experiment U8:
Changed variables compared to baseline condition:
belief (Harry, facilitate (avoid, Barry, reduce_anger_self, -1))
has_ambition(Harry, reduce_anger_self, 1)

	Harry	Barry	Gary
Harry		Involvement = 0.22 Distance = 0 Use Intentions = 0 Expected Satisfaction = 0.132 EU of avoid =-1 A-->I= 0.5 A-->D= -0.5 ES of PA = 0.488 ES of NA = 0.312 ES of CH = 0.088 ES of AV = 0.19 Performed action = **comfort**	Involvement = 0.12 Distance = 0.1 Use Intentions = 0 Expected Satisfaction = 0.092 EU of all actions = 0 A-->I=0 A-->D=0
Barry	Involvement = 0.12 Distance = 0.1 Use Intentions = 0 Expected Satisfaction = 0.092 EU of action = 0 A-->I=0 A-->D=0 ES of PA = 0.408 ES of NA = 0.392 ES of CH = 0.078 ES of AV = 0.47 Performed action = **avoid**		Involvement = 0.12 Distance = 0.1 Use Intentions = 0 Expected Satisfaction = 0.092 EU of actions = 0 A-->I=0 A-->D=0 ES of PA = 0.408 ES of NA = 0.392 ES of CH = 0.078 ES of AV = 0.47 Performed action = **avoid**
Gary	Involvement = 0.12 Distance = 0.1 Use Intentions = 0 Expected Satisfaction = 0.092 EU of action = 0 A-->I=0 A-->D=0 ES of PA = 0.408 ES of NA = 0.392 ES of CH = 0.078 ES of AV = 0.47 Performed action = **avoid**	Involvement = 0.12 Distance = 0.1 Use Intentions = 0 Expected Satisfaction = 0.092 EU of action = 0 A-->I=0 A-->D=0 ES of PA = 0.408 ES of NA = 0.392 ES of CH = 0.078 ES of AV = 0.47 Performed action = **avoid**	

Experiment F1:
Changed variables compared to baseline condition:
designed(Harry, beautiful, 1)
designed(Harry, good, 1)
designed(Harry, realistic, 1)

	Harry	Barry	Gary
Harry		Involvement = 0.0745 Distance = 0.1475 Use Intentions = 0 Expected Satisfaction = 0.1034 EU of all actions = 0 ES of PA = 0.3708 ES of NA = 0.4292 ES of CH = 0.07405 ES of AV = 0.507 Performed action = **avoid**	Involvement = 0.0745 Distance = 01475 Use Intentions = 0 Expected Satisfaction = 0.1034 EU of all actions = 0 ES of PA = 0.3708 ES of NA = 0.4292 ES of CH = 0.07405 ES of AV = 0.507 Performed action = **avoid**
Barry	Involvement = 0.4895 Distance = -0.0025 Use Intentions = 0 Expected Satisfaction = 0.2932 EU of action = 0 ES of PA = 0.5968 ES of NA = 0.2032 ES of CH = 0.19505 ES of AV = 0.2545 Performed action = **comfort**		Involvement = 0.12 Distance = 0.1 Use Intentions = 0 Expected Satisfaction = 0.092 EU of all actions = 0
Gary	Involvement = 0.4895 Distance = -0.0025 Use Intentions = 0 Expected Satisfaction = 0.2932 EU of action = 0 ES of PA = 0.5986 ES of NA = 0.2032 ES of CH = 0.19505 ES of AV = 0.2545 Performed action = **comfort**	Involvement = 0.12 Distance = 0.1 Use Intentions = 0 Expected Satisfaction = 0.092 EU of all actions = 0 ES of all actions = Performed action =	

Experiment F2:
Changed variables compared to baseline condition:
designed(Gary, ugly, 1)
designed(Gary, bad, 1)
designed(Gary, unrealistic, 1)

	Harry	Barry	Gary
Harry		Involvement = 0.12 Distance = 0.1 Use Intentions = 0 Expected Satisfaction = 0.092 EU of all actions = 0	Involvement = 0.1105 Distance = 0.6275 Use Intentions = 0 Expected Satisfaction = 0.3986 EU of all actions = 0 ES of PA = 0.1932 ES of NA = 0.6068 ES of CH = 0.23245 ES of AV = 0.633 Performed action = **avoid**
Barry	Involvement = 0.12 Distance = 0.1 Use Intentions = 0 Expected Satisfaction = 0.092 EU of action = 0		Involvement = 0.1105 Distance = 0.6275 Use Intentions = 0 Expected Satisfaction = 0.3986 EU of all actions = 0 ES of PA = 0.1932 ES of NA = 0.6068 ES of CH = 0.23245 ES of AV = 0.633 Performed action = **avoid**
Gary	Involvement = 0.0855 Distance = 0.1525 Use Intentions = 0 Expected Satisfaction = 0.1086 EU of action = 0 ES of PA = 0.3732 ES of NA = 0.4268 ES of CH = 0.07995 ES of AV = 0.503 Performed action = **avoid**	Involvement = 0.0855 Distance = 0.1525 Use Intentions = 0 Expected Satisfaction = 0.1086 EU of all actions = 0 ES of PA = 0.3732 ES of NA = 0.4268 ES of CH = 0.07995 ES of AV = 0.503 Performed action = **avoid**	

Experiment F3:

Changed variables compared to baseline condition:

designed(Harry, beautiful, 1) designed(Harry, good, 1)
designed(Harry, realistic, 1) designed(Gary, ugly, 1)
designed(Gary, bad, 1) designed(Gary, unrealistic, 1)

	Harry	**Barry**	**Gary**
Harry		Involvement = 0.0745 Distance = 0.1475 Use Intentions = 0 Expected Satisfaction = 0.1034 EU of all actions = 0	Involvement = 0.0650 Distance = 0.675 Use Intentions = 0 Expected Satisfaction = 0.4180 EU of all actions = 0 ES of PA = 0.156 ES of NA = 0.644 ES of CH = 0.2285 ES of AV = 0.67 Performed action = **avoid**
Barry	Involvement = 0.4895 Distance = -0.0025 Use Intentions = 0 Expected Satisfaction = 0.2932 EU of action = 0		Involvement = 0.1105 Distance = 0.6275 Use Intentions = 0 Expected Satisfaction = 0.3986 EU of all actions = 0 ES of PA = 0.1932 ES of NA = 0.6068 ES of CH = 0.23245 ES of AV = 0.633 Performed action = **avoid**
Gary	Involvement = 0.4550 Distance = 0.0500 Use Intentions = 0 Expected Satisfaction = 0.2830 EU of action = 0 ES of PA = 0.562 ES of NA = 0.238 ES of CH = 0.197 ES of AV = 0.2875 Performed action = **comfort**	Involvement = 0.0855 Distance = 0.1525 Use Intentions = 0 Expected Satisfaction = 0.1086 EU of all actions = 0	

Experiment F4:
Changed variables compared to baseline condition:
designed(Gary, beautiful, 1)
designed(Gary, realistic, 1)
designed(Gary, bad, 1)
designed(Gary, good, 1)
designed(Gary, ugly, 1)
designed(Gary, unrealistic, 1)

	Harry	Barry	Gary
Harry		Involvement = 0.12 Distance = 0.1 Use Intentions = 0 Expected Satisfaction = 0.092 EU of all actions = 0	Involvement = 0.48 Distance = 0.525 Use Intentions = 0 Expected Satisfaction = 0.411 EU of all actions = 0 ES of PA = 0.382 ES of NA = 0.418 ES of CH = 0.3495 ES of AV = 0.4175 Performed action = **fight**
Barry	Involvement = 0.12 Distance = 0.1 Use Intentions = 0 Expected Satisfaction = 0.092 EU of action = 0		Involvement = 0.48 Distance = 0.525 Use Intentions = 0 Expected Satisfaction = 0.411 EU of all actions = 0 ES of PA = 0.382 ES of NA = 0.418 ES of CH = 0.3495 ES of AV = 0.4175 Performed action = **fight**
Gary	Involvement = 0.04 Distance = 0.2 Use Intentions = 0 Expected Satisfaction = 0.128 EU of action = 0 ES of PA = 0.336 ES of NA = 0.464 ES of CH = 0.076 ES of AV = 0.54 Performed action = **avoid**	Involvement = 0.04 Distance = 0.2 Use Intentions = 0 Expected Satisfaction = 0.128 EU of all actions = 0 ES of PA = 0.336 ES of NA = 0.464 ES of CH = 0.076 ES of AV = 0.54 Performed action = **avoid**	

Experiment C1:

Changed variables compared to baseline condition:

designed(Gary, beautiful, 1) designed(Gary, good, 1)
designed(Gary, realistic, 1) designed(Gary, ugly, 1)
designed(Gary, bad, 1) designed(Gary, unrealistic, 1)
belief (Harry, facilitate (criticize, Gary, reduce_anger_self, 1))
has_ambition(Harry, reduce_anger_self, 1)

	Harry	Barry	Gary
Harry		Involvement = 0.12 Distance = 0.1 Use Intentions = 0 Expected Satisfaction = 0.092 EU of all actions = 0	A-->I = 0.5 A-->D = 0.5 Involvement = 0.58 Distance = 0.625 Use Intentions = 1 Expected Satisfaction = 0.691 EU of criticize = 1 ES of PA = 0.382 ES of NA = 0.418 ES of CH = 0.7195 ES of AV = 0.3975 Performed action = **criticize**
Barry	Involvement = 0.12 Distance = 0.1 Use Intentions = 0 Expected Satisfaction = 0.092 EU of action = 0		Involvement = 0.48 Distance = 0.525 Use Intentions = 0 Expected Satisfaction = 0.411 EU of all actions = 0 ES of PA = 0.382 ES of NA = 0.418 ES of CH = 0.3495 ES of AV = 0.4175 Performed action = **fight**
Gary	Involvement = 0.04 Distance = 0.2 Use Intentions = 0 Expected Satisfaction = 0.128 EU of action = 0 ES of PA = 0.336 ES of NA = 0.464 ES of CH = 0.076 ES of AV = 0.54 Performed action = **avoid**	Involvement = 0.04 Distance = 0.2 Use Intentions = 0 Expected Satisfaction = 0.128 EU of all actions = 0 ES of PA = 0.336 ES of NA = 0.464 ES of CH = 0.076 ES of AV = 0.54 Performed action = **avoid**	

Experiment C2:
Changed variables compared to baseline condition:
designed(Gary, ugly, 1) designed(Gary, bad, 1)
designed(Gary, unrealistic, 1)
belief (Harry, facilitate (comfort, Gary, reduce_anger_self, 1))
has_ambition(Harry, reduce_anger_self, 1)

	Harry	Barry	Gary
Harry		Involvement = 0.12 Distance = 0.1 Use Intentions = 0 Expected Satisfaction = 0.092 EU of all actions = 0	A-->I = 0.75 A-->D = -0.75 Involvement = 0.1105 Distance = 0.4775 Use Intentions = 1 Expected Satisfaction = 0.5386 EU of all actions = 0 ES of PA = 0.5132 ES of NA = 0.4868 ES of CH = 0.24745 ES of AV = 0.513 Performed action = **comfort**
Barry	Involvement = 0.12 Distance = 0.1 Use Intentions = 0 Expected Satisfaction = 0.092 EU of action = 0		Involvement = 0.1105 Distance = 0.6275 Use Intentions = 0 Expected Satisfaction = 0.3986 EU of all actions = 0 ES of PA = 0.1932 ES of NA = 0.6068 ES of CH = 0.23245 ES of AV = 0.633 Performed action = **avoid**
Gary	Involvement = 0.0855 Distance = 0.1525 Use Intentions = 0 Expected Satisfaction = 0.1086 EU of action = 0 ES of PA = 0.3732 ES of NA = 0.4268 ES of CH = 0.07995 ES of AV = 0.503 Performed action = **avoid**	Involvement = 0.0855 Distance = 0.1525 Use Intentions = 0 Expected Satisfaction = 0.1086 EU of all actions = 0 ES of PA = 0.3732 ES of NA = 0.4268 ES of CH = 0.07995 ES of AV = 0.503 Performed action = **avoid**	

Experiment C3:
Changed variables compared to baseline condition:
designed(Gary, beautiful, 1) designed(Gary, good, 1)
designed(Gary, realistic, 1)
belief (Harry, facilitate (fight, Gary, reduce_anger_self, 1))
has_ambition(Harry, reduce_anger_self, 1)

	Harry	**Barry**	**Gary**
Harry		Involvement = 0.12 Distance = 0.1 Use Intentions = 0 Expected Satisfaction = 0.092 EU of all actions = 0	A-->I = 0.25 A-->Dt = 0.75 Involvement = 0.5395 Distance = 0.1475 Use Intentions = 1 Expected Satisfaction = 0.553 EU of all actions = 0 ES of PA = 0.5568 ES of NA = 0.4432 ES of CH = 0.26005 ES of AV = 0.2745 Performed action = **Comfort**
Barry	Involvement = 0.12 Distance = 0.1 Use Intentions = 0 Expected Satisfaction = 0.092 EU of action = 0		Involvement = 0.4895 Distance = -0.0025 Use Intentions = 0 Expected Satisfaction = 0.293 EU of all actions = 0 ES of PA = 0.5986 ES of NA = 0.2032 ES of CH = 0.19505 ES of AV = 0.2545 Performed action = Comfort
Gary	Involvement = 0.07 Distance = 0.1475 Use Intentions = 0 Expected Satisfaction = 0.103 EU of action = 0 ES of PA = 0.3708 ES of NA = 0.4292 ES of CH = 0.07405 ES of AV = 0.507 Performed action = **Avoid**	Involvement = 0.0745 Distance = 0.1475 Use Intentions = 0 Expected Satisfaction = 0.103 EU of action = 0 ES of PA = 0.3708 ES of NA = 0.4292 ES of CH = 0.07405 ES of AV = 0.507 Performed action = **Avoid**	

Experiment C4:

Changed variables compared to baseline condition:

designed(Gary, beautiful, 1) designed(Gary, good, 1)
designed(Gary, realistic, 1) designed(Gary, ugly, 1)
designed(Gary, bad, 1) designed(Gary, unrealistic, 1)
belief (Harry, facilitate (avoid, Gary, reduce_anger_self, 1))
has_ambition(Harry, reduce_anger_self, 1)

	Harry	Barry	Gary
Harry		Involvement = 0.12 Distance = 0.1 Use Intentions = 0 Expected Satisfaction = 0.092 EU of all actions = 0	A-->I = -0.5 A-->D = 0.5 Involvement = 0.38 Distance = 0.625 Use Intentions = 1 Expected Satisfaction = 0.651 EU of all actions = 0 ES of PA = 0.302 ES of NA = 0.498 ES of CH = 0.3395 ES of AV = 0.6975 Performed action = **Avoid**
Barry	Involvement = 0.12 Distance = 0.1 Use Intentions = 0 Expected Satisfaction = 0.092 EU of action = 0		Involvement = 0.48 Distance = 0.525 Use Intentions = 0 Expected Satisfaction = 0.411 EU of all actions = 0 ES of PA = 0.382 ES of NA = 0.418 ES of CH = 0.3495 ES of AV = 0.4175 Performed action = **Fight**
Gary	Involvement = 0.04 Distance = 0.2 Use Intentions = 0 Expected Satisfaction = 0.128 EU of action = 0 ES of PA = 0.336 ES of NA = 0.464 ES of CH = 0.076 ES of AV = 0.54 Performed action = **Avoid**	Involvement = 0.04 Distance = 0.2 Use Intentions = 0 Expected Satisfaction = 0.128 EU of all actions = 0 ES of PA = 0.336 ES of NA = 0.464 ES of CH = 0.076 ES of AV = 0.54 Performed action = **Avoid**	

PART II MODELING INVOLVEMENT BETWEEN AGENTS

CHAPTER 4

Affective Agents Perceiving Each Other's Actions

This chapter appeared as Pontier, M., and Siddiqui, G.F., Affective Agents Perceiving each other's Actions. In: Otamendi, J., Bargiela, A., Montes, J.L., Pedrera, L.M.D. (eds.), Proceedings of the 23rd European Conference on Modeling and Simulation, ECMS'2009. European Council for Modeling and Simulation, 2009, pp. 194-202.

Affective Agents Perceiving Each Other's Actions

Matthijs Pontier[1,2] and Ghazanfar F. Siddiqui[1,2]

[1]VU University, Department of Artificial Intelligence
De Boelelaan 1081, 1081 HV Amsterdam, The Netherlands
{mpr210, ghazanfa}@few.vu.nl
[2]VU University, Center for Advanced Media Research Amsterdam
Buitenveldertselaan 3, 1082 VA Amsterdam, The Netherlands

Abstract. In this paper, an extension of a formalized model of affective decision making is presented, based on the informally described I-PEFiC model. This extension manages that the actions agents undertake have an effect on other agents. The agents change their perceptions and beliefs about other agents if actions are taken. Further, the anger level of the agents is simulated. Simulation experiments show that the actions of agents can change the beliefs and the perceptions of another agent so much that the other agent changes its mind and chooses to perform another action than it was currently doing.

Keywords : Agent-Based Modeling, Affective Decision-Making, Perception.

1 Introduction

From the last decade, a lot of research has been dedicated to developing Intelligent Virtual Agents (IVAs) with more realistic graphical representations. However, these agents often do not show very realistic human-like emotional behavior. For example, many IVAs can show emotions by the means of facial expressions or the tone in their voice, but most of them still struggle to show the right emotions at the right moment (e.g., emotion regulation [16]) and moods [5]). Let alone actually understanding and reacting empathically to the emotional state of other agents, or human users. Previous research has shown that closely mimicking humans is important for an agent to increase human involvement in a virtual environment (e.g., [20]).

From Frijda's point of view, emotions are goal driven [10]. The emotional system examines the surroundings for related stimuli that are either beneficial or harmful for the concerns, motives, and goals of the individual [10]. According to broaden-and-build theory, positive emotions are vehicles for individual growth and social connections: by building people's personal and social resources, positive emotions transform people for the better, giving them better lives in the future [9]. Previous research also showed that human beings usually make unconscious rather than conscious decisions [3].

We created agents for imitating human behavior, who can recognize each other as a personal friend as well as means to an end. The Interactive model of Perceiving and Experiencing Fictional Characters (I-PEFiC), which is based on the theory of Frijda [10], was taken as a foundation with regard to recognizing each other as a personal friend [19]. Within this model an agent A can compute the trade-off between how involved it is with another agent (e. g., Agent B is good) and what keeps the agent at a distance (e. g., Agent B is bad) [6]. This involvement-distance trade-off is the outcome of assessing the features of an agent on several dimensions. Use intentions are calculated additionally that prompt the agent to carry out actions towards another agent. These actions are based on goals, which play a role in the judgment formation of the agent about the other agent, but also more affective influences are taken into account. The previous decision making model [15] only describes the affective decision making process itself. Simulation experiments with this model showed that in situations where this can be considered human-like, agents make affective decisions rather than decisions that would be the best rational decision (i. e., the decision option with the highest expected utility). However, it did not explain the effects of the performed actions on the agents' perceptions, beliefs and levels of anger. Neither did it explain how these changed beliefs and perceptions influenced the following decision making processes.

In this paper we improved the affective decision making model [15], so that the agents update their beliefs and perceptions of the ethics and affordances of other agents when actions are being performed. Further, the effects of these actions on the emotions of the agents are simulated. These changes on their turn influence the decision making process in the agents. This enables the agents to change their mind and decide to change the action they want to perform. The simulation experiments described in this paper will show what kind of effects the actions have on the agents' beliefs, perceptions, and emotions and how this affects their decision making.

2 Modeling Approach

Modeling the various aspects involved in affective decision making in an integrated manner poses some challenges. On the one hand, qualitative aspects have to be addressed, such as performing an action. On the other hand, quantitative aspects have to be addressed, such as levels of anger. The modeling approach based on the modeling language LEADSTO [7] fulfils these needs. It integrates qualitative, logical aspects such as used in approaches based on temporal logic (e. g., [4]) with quantitative, numerical aspects such as used in Dynamical Systems Theory (e. g., [2], [17]).

In LEADSTO, direct temporal dependencies between state properties in two successive states are modeled by executable dynamic properties defined as follows. Let a and b be state properties of the form "conjunction of literals" (where a literal is an atom or the negation of an atom), and e, f, g, h non-negative real numbers. Then in the leads to language $a \twoheadrightarrow_{e, f, g, h} b$, means:

> *If state property* a *holds for a certain time interval with duration* g*, then after some delay (between* e *and* f*) state property* b *will hold for a certain time interval of length* h.

Here, atomic state properties can have a qualitative, logical format, such as an expression desire(d), expressing that desire d occurs, or a quantitative, numerical format such as has_value(x, v) expressing that variable x has value v.

2 Implementation

I-PEFiC is a model (Figure 1) that is empirically well validated (e.g., [19], [20], [21]). The I-PEFiC model has three phases: encoding, comparison, and response [21].
During encoding, the user appraises an agent's features for their level of ethics (good or bad), aesthetics (beautiful or ugly), and epistemics (realistic or unrealistic). During the encoding, moreover, the user evaluates in how far the agent system has affordances (aids or obstacles), which make the agent useful as a computer tool or not.

In the comparison phase, the features are judged for similarity (similar or dissimilar) (e.g., "I am not like the agent"), relevance of features to user goals (relevant or irrelevant) and valence to goals (positive or negative outcome expectancies). The measures in the encode phase - moderated by the factors in the comparison phase - determine the responses, that is, the levels of involvement with and distance towards the embodied agent. Moreover, the intention to use the agent as a tool indicates actual use and together with involvement and distance, this determines the overall satisfaction of the user with the agent; in our case of Agent A with Agent B. The I-PEFiC model has been formalized in [6].

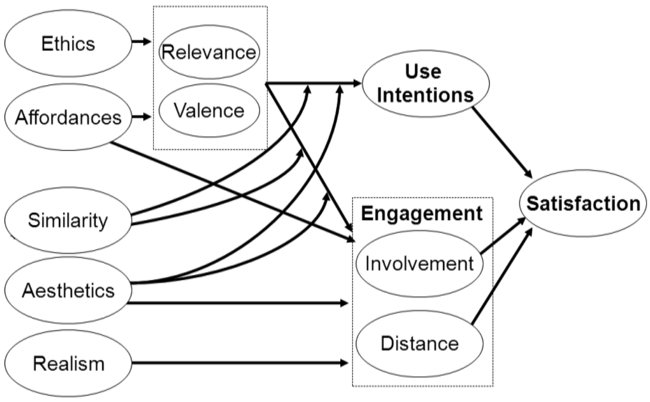


Figure 1. Graphical Representation of I-PEFiC

The model presented in this paper is an extension of a model of affective decision making [15] based on a formalization of the I-PEFiC model [6] in the LEADSTO environment [7]. In this decision making model, decisions are made based on rational as well as affective processing. In the model an agent has desired and undesired goal-states. The agent perceives affordances of the other agents by means of beliefs that the agents facilitate or inhibit reaching certain goal-states. These perceived affordances of other agents are compared to the goal-states it wants to achieve or avoid. While doing this, it can reason about the outcome expectancies of using the other agent for a certain action (e. g., comforting, fighting, or criticizing). In humans, such outcome expectancies lead to certain quick and mostly subconsciously generated action tendencies. In our agents, as in humans, action tendencies influence the experienced involvement and distance towards the other agent. The involvement and distance towards another agent are combined with the use intentions of that agent and the expected utilities of the possible actions to calculate the expected levels of satisfaction of these actions. These expected levels of satisfaction are compared to reach a final decision that is based on rationality as well as affective influences.

We created a library of actions the agents can perform. In this action library, the type of each action is specified. Actions can be specified as

(1) Positive approach
(2) Negative approach
(3) Change
(4) Avoid

In this paper, the action library consists of one action for each type. Comfort is an action of the type positive approach, fight is an action of the type negative approach, criticize is an action of the type change, and avoid obviously is an action of the type avoid. If an agent tries to perform an action of type 1, 2, or 3 towards another agent, while the other agent is avoiding the agent, it will not succeed in performing this action. If an agent performs an action towards another agent, this affects the agent that is the object of the action, as well as the agent that is performing the action itself. The formulas and values not described in this paper can be found in [24].

Adjusting the perceived ethics

If an agent performs an action towards another agent, the agent that is the object of the action can change its perception of the goodness and badness of the agent performing the action. For example, if an agent fights another agent, the agent that is the object of this action will probably decrease its perception of the goodness of the fighting agent, and increase its perception of the badness of the fighting agent. To establish this change in perception, the bias (in the range [0, 2]) changes according to the actions that are being performed. A bias > 1 means overestimation, and a bias < 1 means underestimation. To calculate the effect of an action on the bias for perceiving the goodness of the agent performing the action, we have developed the following formula:

$$\text{new\_bias(good)} = p_{good} * \text{old\_bias(good)} + (1-p_{good}) * v_{(agent,\ action,\ good)}$$

In this formula, new_bias(good) is the new value of the bias, old_bias(good) is the old value of the bias, and the persistency factor p_{good} is the proportion of the old bias that is taken into account to determine the new bias for perceiving goodness of the agent that is performing the action. In this paper, for clarity in the simulation experiments this persistency factor is set to 0.85 for all agents, but this could just as easy be personalized per agent. The new contribution to the bias is $v_{(agent,\ action,\ good)}$, a value that an agent attaches to the goodness of being the object of the performed action. In practice this means that if a certain action is performed towards an agent, the bias of good will move towards the value the agent attaches to being the object of that action. In this paper, all the agents attach the same, arbitrarily chosen values to being the object of actions, but also this could easily be personalized per agent. As the biases are in the range [0, 2], these values are also in the range [0, 2]. The values can be found in Table 1.

Table 1. The Values of Goodness the Agents Attach to Being the Object of an Action

Action	Value
Comfort	1.5
Fight	0.5
Criticize	1.25
Avoid	0.75

The actions for calculating the effect of an action on the bias for perceiving badness are calculated in a similar way as the effect on the bias for perceiving goodness. The only difference is that the values used in this formula are the values the agents attach to the badness of being the object to the performed action. These values can be found in Table 2.

new_bias(bad) =
p_{bad} * old_bias(bad) + (1- p_{bad}) * $v_{agent, action, bad}$

Table 2. The Values of Badness the Agents Attach to Being the Object of an Action

Action	Value
Comfort	0.5
Fight	1.5
Criticize	1.25
Avoid	1.25

Adjusting the perceived affordances

If an agent performs an action towards another agent, the agent that is the object of the action can also change its beliefs about the affordances of the agent performing the action. Beliefs have a value in the domain [-1, 1]. A belief of 1 represents a strong belief that the statement the belief is about is true, and a belief of -1 represents a strong belief that it is not true. For example, if an agent fights another agent, the belief that avoiding the other agent helps to reduce your anger might increase. To calculate the effect of an action on beliefs about the agent performing the action, we have developed the following formulas:

if beliefchange ≥ 0:
new_belief = old_belief + belief_adaptation * beliefchange * ((1 - old_belief) / 2)

if beliefchange < 0:
new_belief = old_belief + belief_adaptation * beliefchange * ((1 + old_belief) / 2)

In these formulas, new_belief is the new value of the belief, old_belief is the old value of the belief, and belief_adaptation is a variable, set at 0.1, that determines the speed with which the beliefs are changed when being the object of an action. The beliefchange is a variable in the range [-1, 1] that determines the change of a belief about the performing agent when an agent is performing an action, or is the object of an action performed by another agent. Multiplying with ((1 - old_belief) / 2) when beliefchange is positive, and with ((1 + old_belief) / 2) when beliefchange is negative manages that the values of the beliefs change less if they approach their boundaries, and prevents them from going out of the domain [-1, 1]. In this paper, the values for beliefchange are the same for all agents, but this could easily be personalized per agent. All beliefchange values can be found in Table 3.

In this table, in the columns the actions that are the cause of the belief change are shown. In the rows the affected beliefs are shown. These beliefs are about an action facilitating a certain goal. For example, if an agent A fights another agent B, agent A will change its belief that comforting agent B will help to reduce his own anger with a beliefchange value of -0.75, as can be seen in Table 3.

Table 3. The beliefchange Values when Actions Are Performed

Affected Belief		**Actions causing the belief change**			
Action	Goal	Comfort	Criticize	Fight	Avoid
Comfort	Self	0.2	-0.2	-0.75	-0.6
Criticize	Self	-0.5	0.9	-0.5	0.6
Fight	Self	-0.9	0.3	0.75	0.4
Avoid	Self	-0.25	0.8	0.9	0.7
Comfort	Others	0.4	0.2	-0.5	-0.3
Criticize	Others	-0.25	-0.25	0.25	-0.4
Fight	Others	-0.8	-0.8	0.25	-0.7
Avoid	Others	-0.35	-0.6	0.75	0.8

Adjusting the emotions of the agents

The actions that an agent performs, and the actions that are performed to an agent, affect the emotions of that agent. The emotion simulated in this paper is the level of anger, but also other or even multiple emotions could be simulated in a similar manner. To calculate the effect of actions performed on the anger level, we have developed the following formula:

$$\text{new\_anger} = p_{anger} / n * \text{old\_anger} + (1 - p_{anger} / n) * \Sigma(\text{changed\_anger}) / n$$

In this formula, new_anger is the new anger level, and old_anger is the old anger level. The persistency factor panger is the proportion of the old anger level that is taken into account to determine the new anger level. In this paper, for clarity in the simulation experiments, this persistency factor is set to 0.95 for all agents, but this could easily be personalized per agent. The number of actions that is taken into account for calculating the new anger level is represented by n. The persistency factor is divided by n, so that there will be less persistency in the anger level if multiple actions are taken into account.

The new contribution to the anger level is the mean of all the changed_anger variables that are attached to the actions taken into account. This changed_anger is a variable that indicates which value the anger level approaches given a certain action. In practice this means that if a certain action is performed towards an agent, the anger level will move towards the value of the changed_anger attached to that action. For example, if an agent fights another agent, this will make its anger level approach 0.7, because if the anger level is very high, fighting another agent might help to release this anger, although it will never help to decrease it to a low anger level. On the other hand, if an agent has a low anger level, fighting another agent will probably increase the level of anger.

In this paper, all the agents attach the same changed_anger values to actions, but it would be just as easy to let each agent have its own personal values. As the anger levels are in the range [0, 1], the changed_anger values are also in the range [0, 1]. The values used for this paper can be found in Table 4.

Table 4. The changed_anger Values the Agents Attach to Specific Actions

Action	Value if subject	Value if object
Comfort	0.2	0.2
Fight	0.7	0.9
Criticize	0.2	0.6
Avoid	0.3	0.7

3 Simulation Experiments

The simulation model introduced in the previous section was used to perform a number of experiments under different parameter settings. In the experiments, the three agents Harry, Barry, and Gary followed a (fictitious) anger management therapy. An infinite number of actions could be inserted in the system, but for clarity, nonetheless, for each action type we inserted only one instance of an action. The action to comfort another agent is a type of positive approach whereas the action for negative approach was to fight another agent. Criticizing another agent was the action associated with change, and the action for avoiding the agent was to simply move away from another agent. In the simulation experiments, a calculation step takes one timepoint and an action takes five timepoints. After these five timepoints the action taken can be reconsidered and changed. The results of the experiment are described below. Notice that in each graph in this paper, the scale on the y-axis can differ.

We expected that the system can generate simulations in which actions lead to changes in perception in such a way, that agents change their mind and perform another action than they would have done before their perception was changed.

Experiment 1: Baseline

To start, an initial experiment was performed that served as a control condition for the rest of the experiments. In this experiment, the designed features for beautiful and ugly (aesthetics), good and bad (ethics), and realistic and unrealistic (epistemics) are all set to 0.5 (see Figure 1). All biases for perceiving these features are set to the neutral value of 1. The agents have no beliefs about other agents (all beliefs are set to 0), and the only ambition they have is to reduce their own anger with an ambition level of 0.5.

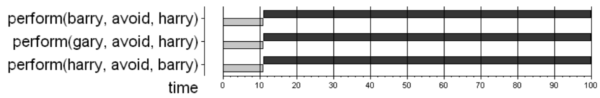


Figure 2. Simulation Results of Experiment 1

In Figure 2, along the X-axis the timepoints are shown, and along the Y-axis several statements are shown. A dark blue bar means the statement holds at that timepoint, and a light blue bar means the statement is false at that timepoint. As can be seen in Figure 2, these settings result in Barry and Harry avoiding Gary, and Harry avoiding Barry after timepoint 11.

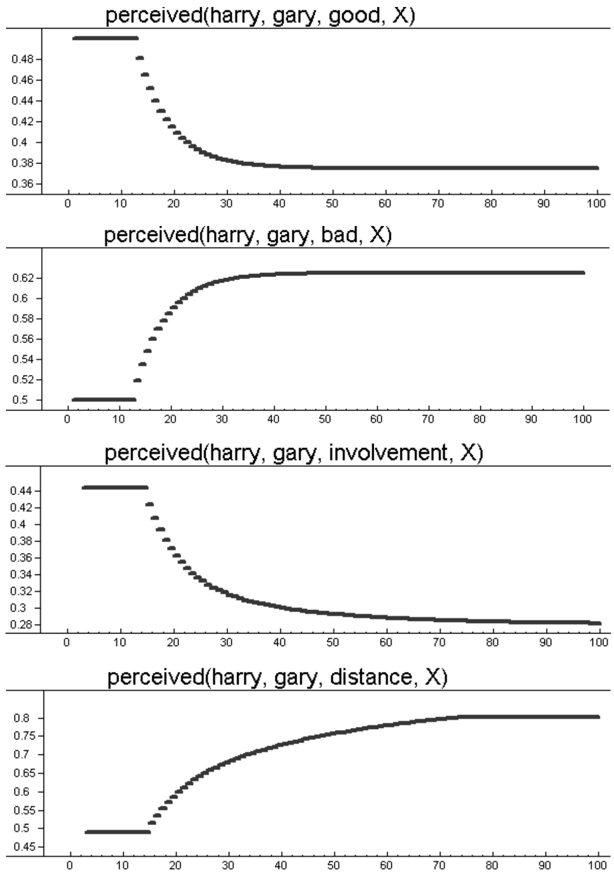


Figure 3. Harry's Involvement, Distance, Perceived Goodness and Perceived Badness of Gary

As can be seen in Figure 3, the agents who are being avoided increase their distance from 0.49 to 0.82, and decrease their involvement from 0.44 to 0.28 towards the avoiding agent. This happens because the perceived goodness of these agents decreases from 0.5 to 0.38 and badness increases from 0.5 to 0.62. Being avoided also causes changes in beliefs.

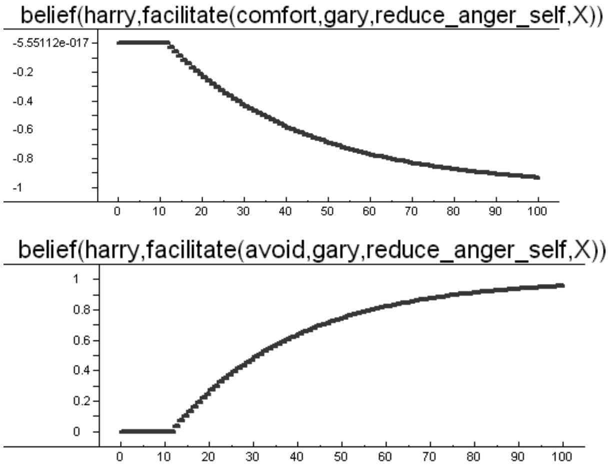


Figure 4. The Beliefs of Barry about Harry during Experiment 1.

As can be seen in Figure 4, the agents who are being avoided start to think that avoiding, criticizing, or fighting the avoiding agent will help them to reduce their own anger, and that comforting the avoiding agent will inhibit their goal of reducing their own anger. As initially the agents have no beliefs about each other, the agents have no intentions to use each other at the start of the simulation.

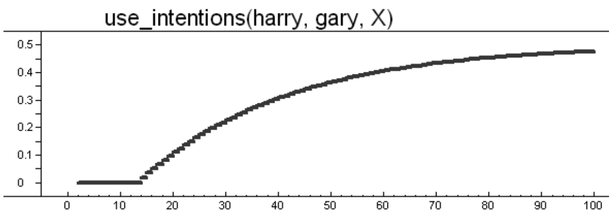


Figure 5. Harry's Intentions to Use Gary during Experiment 1.

Because the agents that are being avoided start to have beliefs about the avoiding agents, their intentions to use that agent also increase from 0 to 0.46, as can be seen in Figure 5.

Being avoided also changes the anger level. As can be seen in Figure 6, the anger level of Barry decreases from 0.60 to 0.50, and the anger level of Harry decreases only from 0.60 to 0.57 (notice the differences in scale on the y-axis). Harry's anger level decreases less because he is avoided by both Barry and Gary. Gary's anger level reduces to 0.30, which is even more than that of Barry, because he is not being avoided by any agent. This shows that being avoided by multiple agents has a greater impact than being avoided by only one agent.

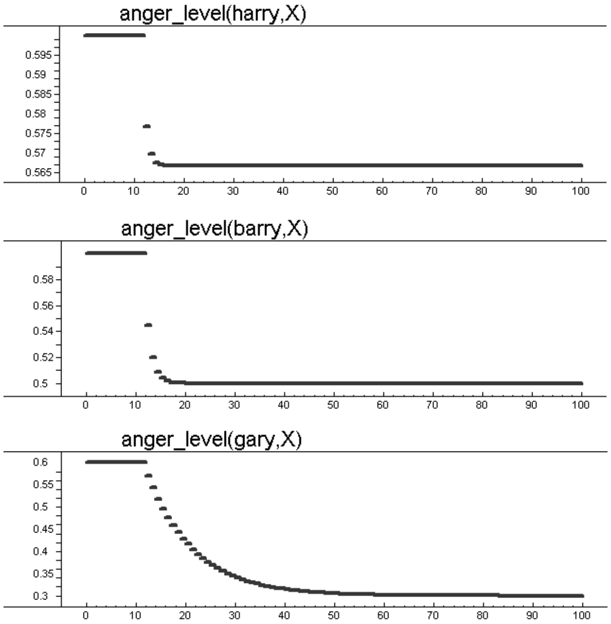


Figure 6. Anger Levels of the Agents in Experiment 1

Experiment 2: Harry beliefs he should not avoid Barry

We performed an experiment in which Harry has a strong belief (value = 1) that avoiding Barry will inhibit his goal of reducing his own anger. Harry also has a stronger ambition to reduce his own anger, with a value of 1 instead of 0.5. The remaining variables have the same values as in the baseline condition.

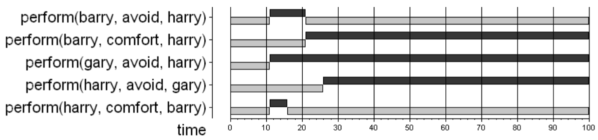


Figure 7. Simulation Results of Experiment 2.

Due to this, instead of avoiding Barry, Harry now tries to comfort Barry at time-point 11, as can be seen in Figure 7. However, he does not succeed in comforting Barry, as Barry is avoiding Harry, just like in the baseline experiment. This causes Harry to stop avoiding Barry at time point 16 and increases his anger level to from 0.53 to 0.70, as can be seen in Figure 8.

Not being comforted anymore also slightly increases Barry's anger level from 0.25 to 0.27 in the five following time steps. In the mean time, Barry has observed that Harry tried to comfort him, which decreases his anger level from 0.60 to 0.25. It also changes his perceptions of the ethics of Harry. As can be seen in Figure 9, Barry starts to see Harry as a good guy, with a perceived goodness of 0.64 and a perceived badness of 0.36.

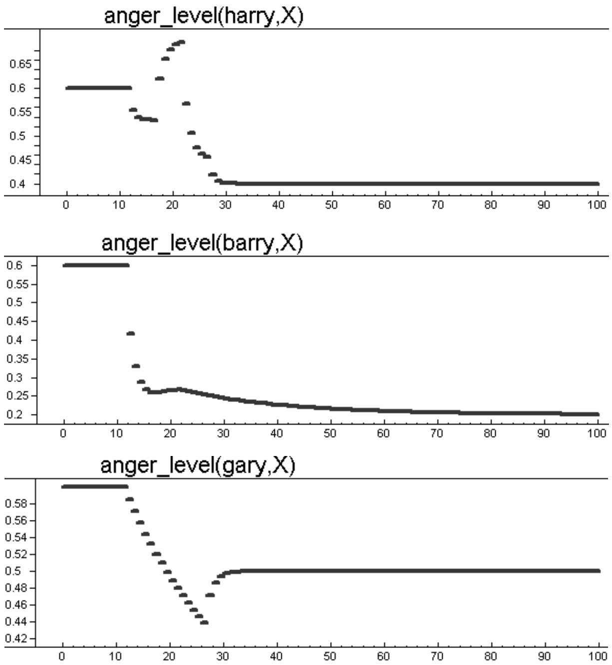


Figure 8. Anger Levels of the Agents in Experiment 2

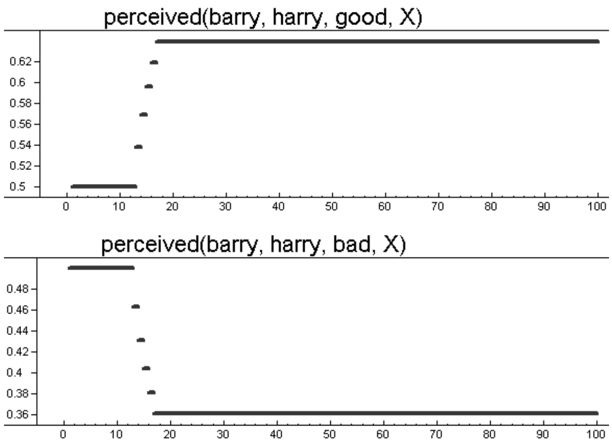


Figure 9. Barry's Perception of the Ethics of Harry during Experiment 2

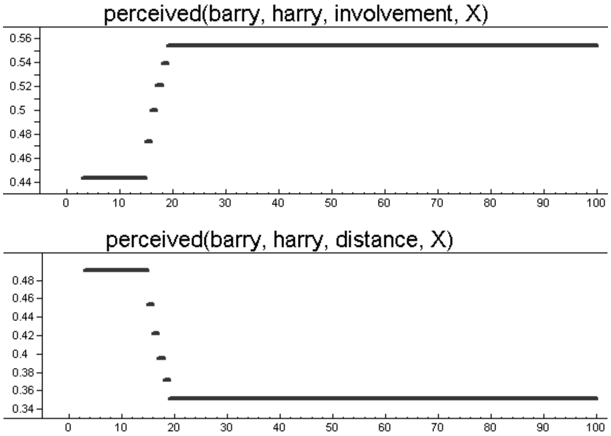


Figure 10. Barry's Involvement and Distance towards Harry during Experiment 2

This also slightly increases Barry's belief that comforting Harry will help him to reduce his anger, while his beliefs that avoiding, criticizing or fighting Barry will help him to reduce his own anger are slightly reduced.

These changes in beliefs and perceptions cause the involvement of Barry towards Harry to increase from 0.44 to 0.56 and the distance from Barry to Harry to decrease from 0.50 to 0.34, as can be seen in Figure 10. This causes him to start comforting Harry instead of avoiding him at timepoint 21.

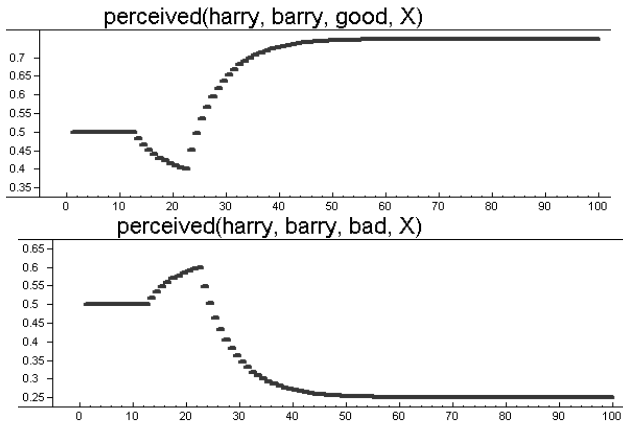


Figure 11. Harry's Perception of the Ethics of Barry during Experiment 2

However, in the mean time, being avoided by Barry has changed Harry's opinion about him. The perception of his goodness has decreased from 0.50 to 0.39, and the perception of his badness has increased from 0.50 to 0.61, as can be seen in Figure 11.

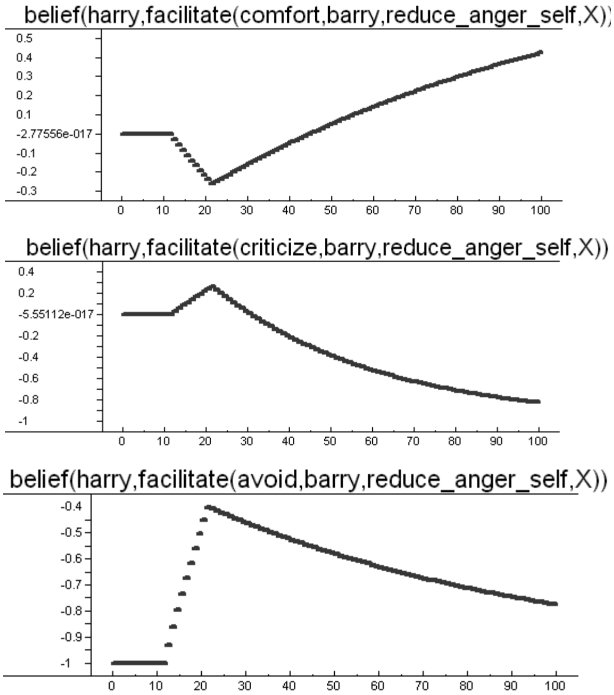


Figure 12. Harry's Beliefs about Barry during Experiment 2

As can be seen in Figure 12, also Harry's beliefs that fighting, criticizing, and especially avoiding Barry will help to reduce his own anger increase, whereas his belief that comforting Barry will help to reduce his own anger decreases. In the mean time, being avoided by Gary has decreased Harry's involvement towards Gary, increased his distance towards Gary, and has increased his beliefs that criticizing, fighting, or avoiding Harry will help him to reduce his own anger in a similar way as in experiment 1.

As can be seen in Figure 13, this increases Harry's expected satisfaction of performing an action towards Gary, namely avoiding him, and this expected satisfaction exceeds the expected satisfaction of performing an action towards Barry. This causes Harry to change his mind and start avoiding Gary at timepoint 26. In the meantime, after Barry started to comfort Harry at timepoint 21, Harry's anger level decreases from 0.70 to 0.40, and Barry's anger level from 0.28 to 0.20.

Also the distance from Harry to Barry decreases from 0.66 to 0.09 and his involvement increases from 0.37 to 0.67, as can be seen in Figure 14. This happens due to an increase in Barry's perception of Harry's goodness from 0.39 to 0.75 and a decrease in perceived badness from 0.61 to 0.25, and because his increasing belief that comforting Harry will help him to reduce his own anger, and fighting, criticizing or avoiding Harry will inhibit his goal to reduce his own anger.

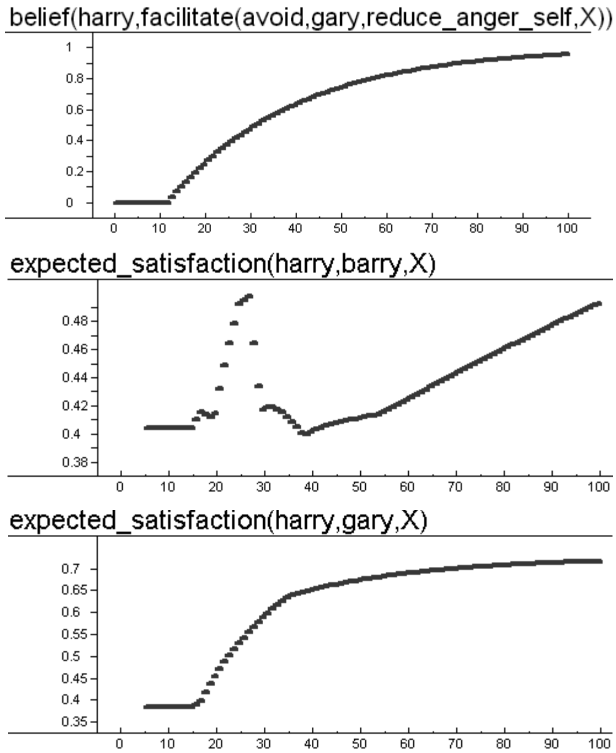


Figure 13: Harry's Belief that Avoiding Gary Will Help Him to Reduce his own Anger, and his Expected Satisfaction of Performing an Action towards Barry and Gary during Experiment 2

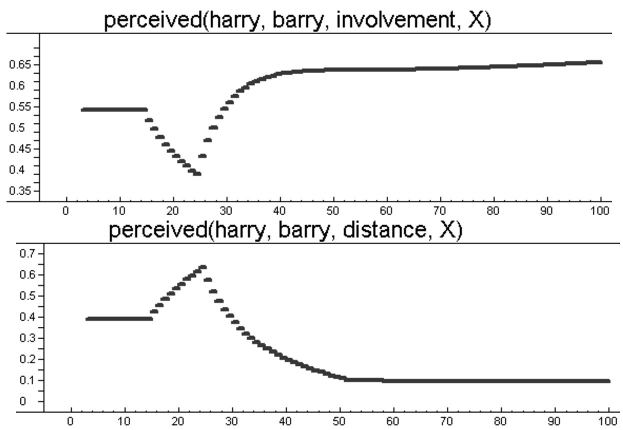


Figure 14: Harry's Involvement and Distance towards Barry during Experiment 2

4 Discussion

In this paper, we presented an extension of a formalized model of affective decision making [15], based on the informally described I-PEFiC model (e.g., [19], [20], [21]). This extension manages that the actions the agents undertake have an effect on the agents. The agents change their perceptions and beliefs about other agents if actions are taken. Further, the anger level of the agents is being simulated. Simulation experiments have been performed to show how the actions affect the agents. Experiment 1 showed that if multiple agents perform an action towards another agent, this has a bigger effect on its anger level than if only one agent would perform that action. Experiment 2 showed that if an agent performs an action towards another agent, this can change the beliefs and the perceptions of the other agents so much that the other agent changes its mind and chooses to perform another action than it was currently doing, leading to a completely different situation than in experiment 1, confirming our expectations as mentioned in section 4. These results are as would have been expected from the theory (e.g., [14], [19], [20]).

In this paper, the simulation experiments are performed in the domain of anger management therapy. The only simulated emotion is anger, and for each type of action there is only one possible action to perform. However, this model could be used for any type of domain, with other, or multiple emotions simulated at the same time. Also as many actions as desired could be added to the action library. This way, the model could be used to perform simulations involving decision-making, emotions, and changing perceptions for any domain.

Of course, a lot of additions could still be made to our model. For instance, the persistency factors for changing the beliefs and perceptions of other agents could be made dependent on the period of time the agents know each other (or on the number of interactions).

Existing models of decision-making usually assume this process to be rational, which would exclude the possibility of emotions playing a role other than disturbing the process [13]. However, humans often make irrational decisions. A good example for this is the Ultimatum game [18], for which behavioral research showed that low offers (20% of total amount) have a 50% chance of being rejected. Participants reported that they found low offers unfair, and therefore out of anger they selected the irrational option [13].

Models of decision making usually have a hedonic bias, and generally try to find the action with the highest expected utility. Some decision theoretic models, such as [11], take emotions into account, but in those models, emotions merely confirm good rational decisions – emotional states as modes of decision making. However, these models cannot explain irrational behavior, where actions with a (relatively) low expected utility are chosen. Our balancing model takes the expected utility as well as involvement-distance trade-offs into account. This way, situations in which emotions overwhelm rationality can be explained and simulated.

There have been a number of approaches to model decision-making based on emotions in autonomous agents. However, none of these studies uses a detailed model of perception of others to explain how these affective influences in the decision making process are generated. Usually, these models somehow assume that emotions are there. For instance, [23] presents a model of emotion-based decision-making, which is an extension of a previous model (e.g., [22]). The model that is presented in this paper assumes a perceptual system, but how this perceptual system actually works is not considered in the paper.

Ahn & Picard [1] present a computational framework of affective-cognitive learning and decision making for affective agents. This model integrates affect and cognition, where 'internal rewards from cognition and emotion' and 'external rewards from the external world' serve as motivations for learning and decision making. In this model emotions are generated based on these rewards, but perceiving others in the world is left out of consideration.

In future research, we plan to combine the model with an existing computational model for emotion regulation [8]. Whereas the current model focuses on the elicitation of emotion, that model addresses the regulation of emotion. Further, we intend to explore where these models and the EMA model [12] complement each other, and use this to further refine the models.

Finally, in a later stage of the project, we will confront our formalization with empirical data of human affective trade-off processes. As soon as the model is validated and adapted, we will start exploring the possibilities to build a robot that can interact with real humans. We hope to develop a robot that can communicate affectively with humans in a more natural way, that is, with a mind of its own, in pursuit of its own goals, and acting emotionally intelligent.

Acknowledgements

We kindly want to thank Tibor Bosse for his input to this paper, and for giving feedback at earlier drafts.

References

1. Ahn, H., and Picard, R. W. 2005. "Affective-cognitive learning and decision making: A motivational reward framework for affective agents." In Tao, J., Tan, T., and Picard, R. W., (Eds.), Affective Computing and Intelligent Interaction, First International Conference, 866–873
2. Ashby, R. 1960. Design for a Brain. Second Edition. Chapman &Hall, London. First edition 1952.
3. Bargh, J. A., and Chartrand, T. L. 1999. "The Unbearable Automaticity of Being." American Psychologist, 54, 462-479.
4. Barringer, H., Fisher, M., Gabbay, D., Owens, R., and Reynolds, M. 1996. The Imperative Future: Principles of Executable Temporal Logic. John Wiley & Sons.
5. Beck, A. T. 1987. "Cognitive models of depression." Journal of Cognitive Psychotherapy, An International Quarterly, 1, 5-37.
6. Bosse, T., Hoorn, J. F., Pontier, M., and Siddiqui, G. F. 2008. "A Robot's Experience of Another Robot: Simulation." In Sloutsky, V., Love, B. C., and McRae, K. (Eds.), Proceedings of the 30th International Annual Conference of the Cognitive Science Society. CogSci'08. 2498-2503.
7. Bosse, T., Jonker, C. M., Meij, L., van der, and Treur, J. 2007. "A Language and Environment for Analysis of Dynamics by Simulation." International Journal of Artificial Intelligence Tools, 16, 435-464.

8. Bosse, T., Pontier, M., and Treur, J. 2007. "A dynamical system modeling approach to Gross' model of emotion regulation." In Lewis, R. L., Polk, T. A., and Laird, J. E. (Eds.), Proceedings of the 8th International Conference on Cognitive Modeling. ICCM'07. Taylor and Francis, 187-192.
9. Fredrickson, B. L. 2001. "The Role of Positive Emotions in Positive Psychology: The broaden-and-build theory of positive emotions." American Psychologist, 56, 218-226.
10. Frijda, N. H. 1986. The emotions. New York: Cambridge University.
11. Gmytrasiewycz, P. J., and Lisetti, C. L. 2001. "Emotions and Personality in Agent Design and Modeling." In Bauer, M., Gmytrasiewicz, P. J., and Vaassileva, J. (Eds.), Lecture Notes in Computer Science, Springer-Verlag Berlin Heidelberg, 237-239.
12. Gratch, J., and Marsella, S. 2004. "A Domain-independent Framework for Modeling Emotion." Journal of Cognitive Systems Research, 5, 4, 269-306.
13. Gutnik, L. A., Hakimzada, A. F., Yoskowitz, N. A., and Patel, V. L. 2006. "The role of emotion in decision-making: A cognitive neuroeconomic approach towards understanding sexual risk behavior." Journal of Biomedical Informetics, 39, 720-736.
14. Hoorn, J. F., Konijn, E. A., Van Vliet, H., and Van der Veer, G. 2007. "Requirements change: Fears dictate the must haves; desires the won't haves" In Journal of Systems and Software, Vol. 80, pp. 328-355.
15. Hoorn, J. F., Pontier, M., and Siddiqui, G. F. 2008. "When the User is Instrumental to Robot Goals: First Try – Agent Uses Agent." In Proceedings of the 2008 IEEE/WIC/ACM International Conference on Intelligent Agent Technology. IAT'08, 296-301.
16. Marsella, S., and Gratch, J. 2003. "Modeling coping behavior in virtual humans: Don't worry, be happy." In Proceedings of Second International Joint Conference on Autonomous Agents and Multiagent Systems. AAMAS'03. ACM Press, 313-320.
17. Port, R. F., and Gelder, T. van (Eds.). 1995. Mind as Motion: Explorations in the Dynamics of Cognition. MIT Press, Cambridge, Mass.
18. Thaler, R. H. 1988. "The Ultimatum Game." Journal of Economic Perspectives, 2, 195-206.
19. Van Vugt, H. C. 2008. Embodied Agents from a User's Perspective. Doctoral dissertation, VU University, Amsterdam.
20. Van Vugt, H. C., Hoorn, J. F., Konijn, E. A., and De Bie Dimitriadou, A. 2006. "Affective affordances: Improving interface character engagement through interaction." International Journal of Human-Computer Studies, 64, 874-888.
21. Van Vugt, H. C., Konijn, E. A., Hoorn, J. F., Eliëns, A., and Keur, I. 2007. "Realism is not all! User engagement with task-related interface characters." Interacting with Computers, 19, 267-280.
22. Velásquez, J. D., 1997. "Modeling Emotions and Other Motivations in Synthetic Agents." In Proceedings of the Fourteenth National Conference on Artificial Intelligence, 10-15
23. Velásquez, J. D., 1998. "Modeling Emotion-Based Decision Making." In Cañamero, D. (Eds.): Emotional and Intelligent: The Tangled Knot of Cognition, 164-169
24. See appendix at the end of the chapter (http://www.few.vu.nl/~mpontier/ECMS-2009-Appendices.pdf).

BIOGRAPHIES

MATTHIJS PONTIER was born in Amsterdam, The Netherlands, and went to the VU University of Amsterdam. He obtained his Bachelor degree for Artificial Intelligence as well as for (Cognitive) Psychology, and got his Masters degree in the direction of Cognitive Science in 2007. Currently, he is employed as a Ph.D. student at the VU University of Amsterdam, both at the department of Artificial Intelligence and the Center for Advanced Media Research Amsterdam (CAMeRA). His e-mail address is: mpontier@few.vu.nl and his website can be found at http://www.few.vu.nl/~mpontier

GHAZANFAR FAROOQ SIDDIQUI was born in Mandi Bahaulddin, Pakistan. He did his Master Degree in Computer Science at Quaid-i-Azam University, Islamabad, Pakistan. He spent seven years teaching graduate and under-graduate courses at University of Agriculture, Faisalabad and Quaid-i-Azam University, Islamabad. Since 2006, he is a PhD student at VU University of Amsterdam, both at the department of Artificial Intelligence and the Center for Advanced Media Research Amsterdam (CAMeRA). His e-mail address is: ghazanfa@few.vu.nl and his website can be found at http://www.few.vu.nl/~ghazanfa

Appendix A: Formulas used in the model

Variable	Meaning	Range
abs(X)	The absolute value of variable or formula X	
max(X, Y)	The minimum value of variables or formulas X and Y	
$A \rightarrow I$	Effect of affordances on involvement	
$A \rightarrow D$	Effect of affordances on Distance	
$\text{Perceived}_{(<Feature>, A1, A2)}$	New value Agent1 perceives of a certain feature of Agent2	[0, 1]
$\text{Designed}_{(<Feature>, A2)}$	Value assigned by 'the designer' to a certain feature of Agent2	[0, 1]
$\text{Bias}_{(A1, A2, feature)}$	Bias that Agent1 has of Agent2 regarding a certain feature	[0, 2]
Inv_Dist_trade_off	Outcome of involvement distance trade off	[0, 1]
γ_{factor}	Variable that determines outcome of trade off for a certain factor	[0, 1]
$\beta_{factor \leftarrow factor}$	(personal) regression weight a factor has for another factor for a certain agent	[0, 1]

To increase readability, the names of the beta weights use the following acronyms:

Variable	Acronym	Variable	Acronym
Good	good	Similarity	sim
Bad	bad	Dissimilarity	ds
Beautiful	bea	Relevance	rel
Ugly	ugly	Irrelevance	irr
Realistic	real	Positive Valence	pv
Unrealistic	unr	Negative Valence	nv
Aid	aid	Involvement	inv
Obstacle	obst	Distance	dis

$\text{Perceived}_{(Beautiful, A1, A2)} = \text{Bias}_{(A1, A2, Beautiful)} * \text{Designed}_{(Beautiful, A2)}$
$\text{Perceived}_{(Ugly, A1, A2)} = \text{Bias}_{(A1, A2, Ugly)} * \text{Designed}_{(Ugly, A2)}$

$\text{Perceived}_{(Realistic, A1, A2)} = \text{Bias}_{(A1, A2,Realistic)} * \text{Designed}_{(Realistic, A2)}$
$\text{Perceived}_{(Unrealistic, A1, A2)} = \text{Bias}_{(A1, A2, Unrealistic)} * \text{Designed}_{(Unrealistic, A2)}$

$\text{Perceived}_{(Good, A1, A2)} = \text{Bias}_{(A1, A2,Good)} * \text{Designed}_{(Good, A2)}$
$\text{Perceived}_{(Bad, A1, A2)} = \text{Bias}_{(A1, A2,Good)} * \text{Designed}_{(Good, A2)}$

$\text{ExpectedUtility}_{(Action, Agent, Goal)} = \text{Belief}_{(facilitates(Action, Agent, Goal))} * \text{Ambition}_{(Goal)}$
$\text{ExpectedUtility}_{(Action, Agent)} = \Sigma(\text{ExpectedUtility}_{(Action, Agent, Goal)})$

$AT_{(Action, Agent)} = \text{ExpectedUtility}_{(Action, Agent)}$

$A \rightarrow I = \beta_{I \leftarrow NA} * AT_{neg\_appr} + \beta_{I \leftarrow PA} * AT_{pos\_appr} + \beta_{I \leftarrow CH} * AT_{change} + \beta_{I \leftarrow AV} * AT_{avoid}$
$A \rightarrow D = \beta_{D \leftarrow NA} * AT_{neg\_appr} + \beta_{D \leftarrow PA} * AT_{pos\_appr} + \beta_{D \leftarrow CH} * AT_{change} + \beta_{D \leftarrow AV} * AT_{avoid}$

$\text{Similarity}_{(A1, A2)} =$
$1- ($
$\beta_{sim \leftarrow good} * \text{abs}(\text{Perceived}_{(Good, A1, A2)} - \text{Perceived}_{(Good, A1, A1)}) +$
$\beta_{sim \leftarrow bad} * \text{abs}(\text{Perceived}_{(Bad, A1, A2)} - \text{Perceived}_{(Bad, A1, A1)}) +$
$\beta_{sim \leftarrow bea} * \text{abs}(\text{Perceived}_{(Beautiful, A1, A2)} - \text{Perceived}_{(Beautiful, A1, A1)}) +$
$\beta_{sim \leftarrow ugly} * \text{abs}(\text{Perceived}_{(Ugly, A1, A2)} - \text{Perceived}_{(Ugly, A1, A1)}) +$
$\beta_{sim \leftarrow real} * \text{abs}(\text{Perceived}_{(Realistic, A1, A2)} - \text{Perceived}_{(Realistic, A1, A1)}) +$
$\beta_{sim \leftarrow unr} * \text{abs}(\text{Perceived}_{(Unrealistic, A1, A2)} - \text{Perceived}_{(Unrealistic, A1, A1)})$

$$
\begin{aligned}
&\text{Dissimilarity}_{(A1,\ A2)} = \\
&\quad \beta_{ds \leftarrow good} * \text{abs}(\text{Perceived}_{(Good,\ A1,\ A2)} - \text{Perceived}_{(Good,\ A1,\ A1)}) + \\
&\quad \beta_{ds \leftarrow bad} * \text{abs}(\text{Perceived}_{(Bad,\ A1,\ A2)} - \text{Perceived}_{(Bad,\ A1,\ A1)}) + \\
&\quad \beta_{ds \leftarrow bea} * \text{abs}(\text{Perceived}_{(Beautiful,\ A1,\ A2)} - \text{Perceived}_{(Beautiful,\ A1,\ A1)}) + \\
&\quad \beta_{ds \leftarrow ugly} * \text{abs}(\text{Perceived}_{(Ugly,\ A1,\ A2)} - \text{Perceived}_{(Ugly,\ A1,\ A1)}) + \\
&\quad \beta_{ds \leftarrow real} * \text{abs}(\text{Perceived}_{(Realistic,\ A1,\ A2)} - \text{Perceived}_{(Realistic,\ A1,\ A1)}) + \\
&\quad \beta_{ds \leftarrow unr} * \text{abs}(\text{Perceived}_{(Unrealistic,\ A1,\ A2)} - \text{Perceived}_{(Unrealistic,\ A1,\ A1)})
\end{aligned}
$$

$$
\begin{aligned}
&\text{Relevance}_{(A1,\ A2)} = \\
&\quad \beta_{rell \leftarrow good} * \text{Perceived}_{(Good,\ A1,\ A2)} + \qquad \beta_{rel \leftarrow bad} * \text{Perceived}_{(Bad,\ A1,\ A2)}
\end{aligned}
$$

$$
\begin{aligned}
&\text{Irrelevance}_{(A1,\ A2)} = 1 + (\\
&\quad \beta_{irr \leftarrow good} * \text{Perceived}_{(Good,\ A1,\ A2)} + \qquad \beta_{irr \leftarrow bad} * \text{Perceived}_{(Bad,\ A1,\ A2)})
\end{aligned}
$$

$$
\begin{aligned}
&\text{Positive\_Valence}_{(A1,\ A2)} = 0.5 + \\
&\quad {}_{pv \leftarrow good} * \text{Perceived}_{(Good,\ A1,\ A2)} + \ {}_{pv \leftarrow bad} * \text{Perceived}_{(Bad,\ A1,\ A2)}
\end{aligned}
$$

$$
\begin{aligned}
&\text{Negative\_Valence}_{(A1,\ A2)} = 0.5 + \\
&\quad \beta_{nv \leftarrow good} * \text{Perceived}_{(Good,\ A1,\ A2)} + \qquad \beta_{nv \leftarrow bad} * \text{Perceived}_{(Bad,\ A1,\ A2)}
\end{aligned}
$$

$$
\begin{aligned}
&\text{Involvement}_{(A1,\ A2)} = 0.25 + \\
&\quad \beta_{inv \leftarrow bea} * \text{Perceived}_{(Beautiful,\ A1,\ A2)} + \\
&\quad \beta_{inv \leftarrow ugly} * \text{Perceived}_{(Ugly,\ A1,\ A2)} + \\
&\quad \beta_{inv \leftarrow real} * \text{Perceived}_{(Realistic,\ A1,\ A2)} + \\
&\quad \beta_{inv \leftarrow unr} * \text{Perceived}_{(Unrealistic,\ A1,\ A2)} + \\
&\quad \beta_{inv \leftarrow pv} * \text{Pos\_Valence}_{(A1,\ A2)} + \\
&\quad \beta_{inv \leftarrow nv} * \text{Neg\_Valence}_{(A1,\ A2)} + \\
&\quad \beta_{inv \leftarrow ps} * \text{Pos\_Valence}_{(A1,\ A2)} * \text{Similarity}_{(A1,\ A2)} + \\
&\quad \beta_{inv \leftarrow ns} * \text{Neg\_Valence}_{(A1,\ A2)} * \text{Similarity}_{(A1,\ A2)} + \\
&\quad \beta_{inv \leftarrow pd} * \text{Pos\_Valence}_{(A1,\ A2)} * \text{Dissimilarity}_{(A1,\ A2)} + \\
&\quad \beta_{inv \leftarrow nd} * \text{Neg\_Valence}_{(A1,\ A2)} * \text{Dissimilarity}_{(A1,\ A2)} + \\
&\quad \beta_{inv \leftarrow pb} * \text{Pos\_Valence}_{(A1,\ A2)} * \text{Perceived}_{(Beautiful,\ A1,\ A2)} + \\
&\quad \beta_{inv \leftarrow nb} * \text{Neg\_Valence}_{(A1,\ A2)} * \text{Perceived}_{(Beautiful,\ A1,\ A2)} + \\
&\quad \beta_{inv \leftarrow pu} * \text{Pos\_Valence}_{(A1,\ A2)} * \text{Perceived}(_{Ugly,\ A1,\ A2)} + \\
&\quad \beta_{inv \leftarrow nu} * \text{Neg\_Valence}_{(A1,\ A2)} * \text{Perceived}(_{Ugly,\ A1,\ A2)} + \\
&\quad \beta_{inv \leftarrow rel} * \text{Relevance}_{(A1,\ A2)} + \\
&\quad \beta_{inv \leftarrow irr} * \text{Irrelevance}_{(A1,\ A2)} + \\
&\quad \beta_{inv \leftarrow rs} * \text{Relevance}_{(A1,\ A2)} * \text{Similarity}_{(A1,\ A2)} + \\
&\quad \beta_{inv \leftarrow is} * \text{Irrelevance}_{(A1,\ A2)} * \text{Similarity}_{(A1,\ A2)} + \\
&\quad \beta_{inv \leftarrow rd} * \text{Relevance}_{(A1,\ A2)} * \text{Dissimilarity}_{(A1,\ A2)} + \\
&\quad \beta_{inv \leftarrow id} * \text{Irrelevance}_{(A1,\ A2)} * \text{Dissimilarity}_{(A1,\ A2)} + \\
&\quad \beta_{inv \leftarrow rb} * \text{Relevance}_{(A1,\ A2)} * \text{Perceived}_{(Beautiful,\ A1,\ A2)} + \\
&\quad \beta_{inv \leftarrow ib} * \text{Irrelevance}_{(A1,\ A2)} * \text{Perceived}_{(Beautiful,\ A1,\ A2)} + \\
&\quad \beta_{inv \leftarrow ru} * \text{Relevance}_{(A1,\ A2)} * \text{Perceived}(_{Ugly,\ A1,\ A2)} + \\
&\quad \beta_{inv \leftarrow iu} * \text{Irrelevance}_{(A1,\ A2)} * \text{Perceived}(_{Ugly,\ A1,\ A2)} \\
&\quad \beta_{inv \leftarrow aff} * A \rightarrow I_{(A1,\ A2)}
\end{aligned}
$$

$Distance_{(A1, A2)} =$

$\beta_{dis \leftarrow bea}$ * Perceived $_{(Beautiful, A1, A2)}$ +
$\beta_{dis \leftarrow ugly}$ * Perceived$_{(Ugly, A1, A2)}$ +
$\beta_{dis \leftarrow real}$ * Perceived $_{(Realistic, A1, A2)}$ +
$\beta_{dis \leftarrow unr}$ * Perceived$_{(Unrealistic, A1, A2)}$ +
$\beta_{dis \leftarrow pv}$ * Pos_Valence $_{(A1, A2)}$ +
$\beta_{dis \leftarrow nv}$ * Neg_Valence $_{(A1, A2)}$ +
$\beta_{dis \leftarrow ps}$ * Pos_Valence $_{(A1, A2)}$ * Similarity$_{(A1, A2)}$ +
$\beta_{dis \leftarrow ns}$ * Neg_Valence $_{(A1, A2)}$ * Similarity$_{(A1, A2)}$ +
$\beta_{dis \leftarrow pd}$ * Pos_Valence $_{(A1, A2)}$ * Dissimilarity$_{(A1, A2)}$ +
$\beta_{dis \leftarrow nd}$ * Neg_Valence $_{(A1, A2)}$ * Dissimilarity$_{(A1, A2)}$ +
$\beta_{dis \leftarrow pb}$ * Pos_Valence $_{(A1, A2)}$ * Perceived $_{(Beautiful, A1, A2)}$ +
$\beta_{dis \leftarrow nb}$ * Neg_Valence $_{(A1, A2)}$ * Perceived $_{(Beautiful, A1, A2)}$ +
$\beta_{dis \leftarrow pu}$ * Pos_Valence $_{(A1, A2)}$ * Perceived$_{(Ugly, A1, A2)}$ +
$\beta_{dis \leftarrow nu}$ * Neg_Valence $_{(A1, A2)}$ * Perceived$_{(Ugly, A1, A2)}$ +
$\beta_{dis \leftarrow rel}$ * Relevance $_{(A1, A2)}$ +
$\beta_{dis \leftarrow irr}$ * Irrelevance $_{(A1, A2)}$ +
$\beta_{dis \leftarrow rs}$ * Relevance $_{(A1, A2)}$ * Similarity$_{(A1, A2)}$ +
$\beta_{dis \leftarrow is}$ * Irrelevance $_{(A1, A2)}$ * Similarity$_{(A1, A2)}$ +
$\beta_{dis \leftarrow rd}$ * Relevance $_{(A1, A2)}$ * Dissimilarity$_{(A1, A2)}$ +
$\beta_{dis \leftarrow id}$ * Irrelevance $_{(A1, A2)}$ * Dissimilarity$_{(A1, A2)}$ +
$\beta_{dis \leftarrow rb}$ * Relevance $_{(A1, A2)}$ * Perceived $_{(Beautiful, A1, A2)}$ +
$\beta_{dis \leftarrow ib}$ * Irrelevance $_{(A1, A2)}$ * Perceived $_{(Beautiful, A1, A2)}$ +
$\beta_{dis \leftarrow ru}$ * Relevance $_{(A1, A2)}$ * Perceived$_{(Ugly, A1, A2)}$ +
$\beta_{dis \leftarrow iu}$ * Irrelevance $_{(A1, A2)}$ * Perceived$_{(Ugly, A1, A2)}$
$\beta_{dis \leftarrow aff}$ * $A \rightarrow D_{(A1, A2)}$

$UseIntentions_{(A1, A2)} = max(ExpectedUtility_{(A1, Action, A2)})$

Involvement-Distance-Tradeoff = γ * max(I, D) + (1-γ) * (I+D) / 2
$\gamma = 0.50$

$Expected\_Satisfaction_{(A1, A2)} = \beta_{ES \leftarrow IDT} * IDT + \beta_{ES \leftarrow UI} * UI$

Expected Satisfaction Positive Approach= $\beta_{ESPA \leftarrow I} * I + \beta_{ESPA \leftarrow D} * (1-D) + \beta_{ESPA \leftarrow EU} * EU_{act}$

Expected Satisfaction Negative Approach = $\beta_{ESNA \leftarrow I} * (1-I) + \beta_{ESNA \leftarrow D} * D + \beta_{ESNA \leftarrow EU} * EU_{act}$

Expected Satisfaction Change = $\beta_{ESCH \leftarrow I} * I + \beta_{ESCH \leftarrow D} * D + \beta_{ESCH \leftarrow EU} * EU_{act}$

Expected Satisfaction Avoid = $\beta_{ESAV \leftarrow I} * (1-I) + \beta_{ESAV \leftarrow D} * D + \beta_{ESAV \leftarrow EU} * EU_{act}$

Appendix B: Parameter settings in the experiments

This appendix contains all parameter settings for the experiments in the paper. These are the parameter settings of the baseline condition. The variables that are changed in other experiments are mentioned in the paper.

pgood = 0.85
pbad = 0.85
panger = 0.95

Table 1: Designed values for the features of each agent

Agent	Feature	Value
Harry	Beautiful	0.5
Harry	Ugly	0.5
Harry	Good	0.5
Harry	Bad	0.5
Harry	Realistic	0.5
Harry	Unrealistic	0.5
Barry	Beautiful	0.5
Barry	Ugly	0.5
Barry	Good	0.5
Barry	Bad	0.5
Barry	Realistic	0.5
Barry	Unrealistic	0.5
Gary	Beautiful	0.5
Gary	Ugly	0.5
Gary	Good	0.5
Gary	Bad	0.5
Gary	Realistic	0.5
Gary	Unrealistic	0.5

Table 2: Biases the agents have in perceiving features of their selves and others

Bias of Agent	For perceiving	Of Agent	Value
Harry	Beautiful	Harry	1
Harry	Beautiful	Barry	1
Harry	Beautiful	Gary	1
Harry	Ugly	Harry	1
Harry	Ugly	Barry	1
Harry	Ugly	Gary	1
Harry	Realistic	Harry	1
Harry	Realistic	Barry	1
Harry	Realistic	Gary	1
Harry	Unrealistic	Harry	1
Harry	Unrealistic	Barry	1
Harry	Unrealistic	Gary	1
Harry	Good	Harry	1
Harry	Good	Barry	1
Harry	Good	Gary	1
Harry	Bad	Harry	1
Harry	Bad	Barry	1
Harry	Bad	Gary	1
Barry	Beautiful	Harry	1
Barry	Beautiful	Barry	1
Barry	Beautiful	Gary	1
Barry	Ugly	Harry	1
Barry	Ugly	Barry	1
Barry	Ugly	Gary	1
Barry	Realistic	Harry	1
Barry	Realistic	Barry	1
Barry	Realistic	Gary	1
Barry	Unrealistic	Harry	1
Barry	Unrealistic	Barry	1
Barry	Unrealistic	Gary	1
Barry	Good	Harry	1
Barry	Good	Barry	1
Barry	Good	Gary	1
Barry	Bad	Harry	1
Barry	Bad	Barry	1
Barry	Bad	Gary	1
Gary	Beautiful	Harry	1
Gary	Beautiful	Barry	1
Gary	Beautiful	Gary	1
Gary	Ugly	Harry	1
Gary	Ugly	Barry	1
Gary	Ugly	Gary	1
Gary	Realistic	Harry	1
Gary	Realistic	Barry	1
Gary	Realistic	Gary	1
Gary	Unrealistic	Harry	1
Gary	Unrealistic	Barry	1
Gary	Unrealistic	Gary	1
Gary	Good	Harry	1
Gary	Good	Barry	1
Gary	Good	Gary	1
Gary	Bad	Harry	1
Gary	Bad	Barry	1
Gary	Bad	Gary	1

Table 3: All values for the regression weights

Weight of X	**on Y**	**Value**
Good	Similarity	0.30
Bad	Similarity	0.20
Beautiful	Similarity	0.20
Ugly	Similarity	0.10
Realistic	Similarity	0.10
Unrealistic	Similarity	0.10
Good	Dissimilarity	0.20.
Bad	Dissimilarity	0.30
Beautiful	Dissimilarity	0.10
Ugly	Dissimilarity	0.20
Realistic	Dissimilarity	0.10
Unrealistic	Dissimilarity	0.10
Good	Relevance	0.70
Bad	Relevance	0.30
Good	Irrelevance	-0.70
Bad	Irrelevance	-0.30
Good	Positive Valence	0.50
Bad	Positive Valence	-0.25
Good	Negative Valence	-0.25
Bad	Negative Valence	0.50
Beautiful	Involvement	0.15
Ugly	Involvement	0.05
Realistic	Involvement	0.10
Unrealistic	Involvement	0.05
Positive Valence	Involvement	0.50
Negative Valence	Involvement	-0.15
Positive Valence * similarity	Involvement	0.10
Negative Valence * similarity	Involvement	-0.15
Positive Valence * dissimilarity	Involvement	-0.10
Negative Valence * dissimilarity	Involvement	0.05
Positive Valence * beautiful	Involvement	-0.05
Negative Valence * beautiful	Involvement	-0.10
Positive Valence * ugly	Involvement	0.05
Negative Valence * ugly	Involvement	-0.05
Relevance	Involvement	0.15
Irrelevance	Involvement	-0.10
Relevance * similarity	Involvement	0.10
Irrelevance * similarity	Involvement	0.02
Relevance * dissimilarity	Involvement	0.03
Irrelevance * dissimilarity	Involvement	-0.02
Relevance * beautiful	Involvement	0.10
Irrelevance * beautiful	Involvement	0.03
Relevance * ugly	Involvement	0.03
Irrelevance * ugly	Involvement	0.01
A-->I	Involvement	0.20
Beautiful	Distance	-0.15
Ugly	Distance	0.20
Realistic	Distance	0.05
Unrealistic	Distance	0.10
Positive Valence	Distance	-0.35
Negative Valence	Distance	0.50
Positive Valence * similarity	Distance	-0.15
Negative Valence * similarity	Distance	0.20
Positive Valence * dissimilarity	Distance	0.07

Negative Valence * dissimilarity	Distance	-0.07
Positive Valence * beautiful	Distance	0.08
Negative Valence * beautiful	Distance	0.15
Positive Valence * ugly	Distance	-0.05
Negative Valence * ugly	Distance	-0.05
Relevance	Distance	0.15
Irrelevance	Distance	-0.05
Relevance * similarity	Distance	-0.10
Irrelevance * similarity	Distance	-0.05
Relevance * dissimilarity	Distance	0.05
Irrelevance * dissimilarity	Distance	0.02
Relevance * beautiful	Distance	-0.10
Irrelevance * beautiful	Distance	-0.05
Relevance * ugly	Distance	0.10
Irrelevance * ugly	Distance	0.05
A-->D	Distance	0.20
Inv_Dist_Tradeoff	Expected Satisfaction	0.80
UseIntentions	Expected Satisfaction	0.20

Table 4: Weights of action tendencies on agent's involvement and distance

Weight	Value	Weight	Value
$\beta_{I \leftarrow PA}$	0.75	$\beta_{D \leftarrow PA}$	-0.75
$\beta_{I \leftarrow NA}$	0.25	$\beta_{D \leftarrow NA}$	0.75
$\beta_{I \leftarrow CH}$	0.50	$\beta_{D \leftarrow CH}$	0.50
$\beta_{I \leftarrow AV}$	-0.50	$\beta_{D \leftarrow AV}$	0.50

Table 5: Values for weights of involvement, distance, and expected utility on the expected satisfaction of performing a type of action

Weight	Value	Weight	Value
$\beta_{ESPA \leftarrow I}$	0.4	$\beta_{ESCH \leftarrow I}$	0.4
$\beta_{ESPA \leftarrow D}$	0.4	$\beta_{ESCH \leftarrow D}$	0.3
$\beta_{ESPA \leftarrow EU}$	0.2	$\beta_{ESCH \leftarrow EU}$	0.3
$\beta_{ESNA \leftarrow I}$	0.4	$\beta_{ESAV \leftarrow I}$	0.5
$\beta_{ESNA \leftarrow D}$	0.4	$\beta_{ESAV \leftarrow D}$	0.3
$\beta_{ESNA \leftarrow EU}$	0.2	$\beta_{ESAV \leftarrow EU}$	0.2

PART III INTEGRATION OF APPRAISAL, INVOLVEMENT AND REGULATION

CHAPTER 5

Comparing Three Computational Models Of Affect

This chapter appeared as Bosse, T., Gratch, J., Hoorn, J.F.,Pontier, M.A., and Siddiqui, G.F., Comparing Three Computational Models of Affect In: Demazeau et al. (eds.), Proceedings of the 8th International Conference on Practical Applications of Agents and Multi-Agent Systems, PAAMS 2010. In Advances in PAAMS, AISC 70, Springer Verlag, 2010, pp. 175-184.

Comparing Three Computational Models of Affect

Tibor Bosse[1], Jonathan Gratch[3], Johan F. Hoorn[2], Matthijs Pontier[1,2] and Ghazanfar F. Siddiqui[1,2,4]

[1]VU University, Department of Artificial Intelligence
De Boelelaan 1081, 1081 HV Amsterdam, The Netherlands
{jfhoorn, mpr210, ghazanfa}@few.vu.nl
[2]VU University, Center for Advanced Media Research Amsterdam
Buitenveldertselaan 3, 1082 VA Amsterdam, The Netherlands
jf.hoorn@camera.vu.nl
[3]University of Southern California, Institute for Creative Technologies
gratch@ict.usc.edu
[4]Quaid-i-Azam University Islamabad, Department of Computer Science,
45320, Pakistan

Abstract. In aiming for behavioral fidelity, artificial intelligence cannot and no longer ignores the formalization of human affect. Affect modeling plays a vital role in faithfully simulating human emotion and in emotionally-evocative technology that aims at being real. This paper offers a short expose about three models concerning the generation and regulation of affect: CoMERG, EMA and I-PEFiCADM, which each in their own right are successfully applied in the agent and robot domain. We argue that the three models partly overlap and where distinct, they complement one another. We provide an analysis of the theoretical concepts, and provide a blueprint of an integration, which should result in a more precise representation of affect simulation in virtual humans.

Keywords: Affect modeling, Cognitive modeling, Virtual agents

1 Introduction

Over the last decade, a virtual explosion can be observed in the amount of novel computational models of affect. Nevertheless, current affect models in software agents are still simplifications compared to human affective complexity. Although many agents currently have the ability to show different emotions by means of facial expressions, it is quite difficult for them to show the right emotion at the right moment. In anticipation of richer interactions between user and agent, this paper explores the possibility to integrate a number of models that are sufficiently similar, while preserving their individual qualities. As a first step into that direction, we compared three models (CoMERG, EMA, and I-PEFiC^{ADM}) of agent affect-generation and affect-regulation (or coping).

We selected three models inspired by some of the most influential theories in the emotion domain to achieve more realistic affective behavior in agents. The theory of Emotion and Adaptation of Smith and Lazarus [12] was formalized by Gratch and Marsella [10] into *EMA*, a model to create agents that demonstrate and cope with (negative) affect. The emotion regulation theory of Gross [5] was used as inspiration by Bosse, Pontier, and Treur [2] to develop *CoMERG* (the Cognitive Model for Emotion

Regulation based on Gross), which simulates the various emotion regulation strategies described by Gross. The concern-driven theory of Frijda [4] was used by Hoorn, Pontier, and Siddiqui [6] to design *I-PEFiC^{ADM}*, a model for building robots that can trade rational for affective choices. We consider these theories because of their adequate mechanisms, simplicity and coherence. Together, they (Frijda, Smith & Lazarus, Gross) cover a large part of emotion theory. All three were inspired by an appraisal model of emotion, which makes them well suited for integration. In addition, these models are already implemented as computational models, which make it easier to integrate them. All three approaches point at important aspects of human affective behavior, but also miss out on something. CoMERG [2] and EMA [10] address the regulation of affective states, but EMA does not regulate positive affect. CoMERG, on the other hand, has no provisions for generating affect, and does not explicitly account for a causal interpretation of the world state. I-PEFiC^{ADM} [6] generates and balances affect but is mute about the different regulation mechanisms. Because the models are complementary to each other, it makes sense to integrate them.

As a first step, the present contribution attempts to align and contrast different affect models as they were derived from the original emotion theories[1]. We will point out what deficiencies should be overcome to build a better artifact for human-agent interaction and to gain more insight into human affective processes.

2 CoMERG, EMA, and I-PEFiC^{ADM}

2.1. CoMERG

According to Gross [5], humans use strategies to influence the level of emotional response to a given type of emotion; for instance, to prevent a person from having a too high or low response level.

In [2], Gross' theory was taken as a basis to develop the emotion regulation model CoMERG. This model, which consists of a set of difference equations combined with logical rules, can be used to simulate the dynamics of the various emotion regulation strategies described by Gross. CoMERG was incorporated into agents in a virtual storytelling application [1]. Following Gross' theory, CoMERG distinguishes five different emotion regulation strategies, which can be applied at different points in the process of emotion generation: *situation selection, situation modification, attentional deployment, cognitive change,* and *response modulation.*

2.2. Emotion & Adaption (EMA) Model

EMA is a computational model of the cognitive antecedents and consequences of emotions posited by appraisal theory, particularly as conceptualized by Smith and Lazarus [12]. A central tenet in cognitive appraisal theories is that appraisal and coping center around a person's *interpretation* of their relationship with the environment. This interpretation is constructed by cognitive processes, summarized by appraisal variables

[1] Note that the presented models embody a particular variant of an affect theory in that they have some unique properties that distinguish them from their original source. Many design choices underlying such models arise from the need to create a working computational system, a challenge the original theorists have never confronted.

and altered by coping responses. To capture this process in computational terms, EMA maintains an explicit symbolic representation of the relationship between events and an agent's internal beliefs, desires and intentions, by building on AI planning to represent the physical relationship between events and their consequences, and BDI frameworks to represent the epistemic factors that underlie human (particularly social) activities.

Appraisal processes characterize this representation in terms of individual appraisal judgments, extending traditional AI concerns with utility and probability:

- Desirability: what is the utility (positive or negative) of the event
- Likelihood: how probable is the outcome of the event.
- Causal attribution: who deserves credit/blame.
- Controllability: can the outcome be altered by actions of the agent.
- Changeability: can the outcome change on its own.

Patterns of appraisal elicit emotional displays, but they also initiate coping processes to regulate the agent's cognitive response to the generated emotion. Coping strategies work in the reverse direction of appraisal, identifying plans, beliefs, desires or intentions to maintain or alter in order to reduce negative emotional appraisals:

- Planning: form an intention to perform some act
- Seek instrumental support: ask someone that controls outcome for help.
- Procrastination: wait for an external event to change the current circumstances.
- Denial: lower the perceived likelihood of an undesirable outcome.
- Mental disengagement: lower utility of desired state.
- Shift blame: shift responsibility for an action toward some other agent.

Strategies give input to the cognitive processes that actually execute these directives. For example, planful coping generates an intention to act, leading a planning system associated with EMA to generate and execute a valid plan to accomplish this act. Alternatively, coping strategies might abandon the goal, lower the goal's importance, or re-assess who is to blame.

EMA is a fully implemented model and has been applied to a number of systems that must simulate realistic human emotional responses. Several empirical studies have demonstrated EMA's effectiveness in modeling emotion [9].

2.3. I-PEFiC^{ADM}

Originally, the empirically validated framework for Perceiving and Experiencing Fictional Characters (PEFiC) described the receiver's reception of literature, theater, and movie characters [7]. Later versions were applied to the embodied-agent domain and supplemented with user interaction possibilities, resulting into the Interactive PEFiC model. I-PEFiC was then used to model affective behavior of robots as a module for Affective Decision Making was added to simulate irrational robot behavior, hence I-PEFiC^{ADM} [7].

The groundwork of I-PEFiC^{ADM} is formed by the cognitive process triplet of an encoding, a comparison, and a response phase. During *encoding*, the robot perceives the user and the situation the user is in. The features of the 'user in a situation' are indexed on four dimensions as a description of what someone is like or does. The robot attributes a level of *ethics* to the user, that is, the robot tries to figure whether the user's character is good or bad. *Aesthetics* is a level of beauty or ugliness that the robot perceives in the user. *Epistemics* is a measure for the realistic or unrealistic representations that the user conveys about him or herself. During the encoding, moreover, the robot looks at the user in terms of *affordances*. Certain aspects of the user may count as helpful or as an obstacle.

In the *comparison* phase, the user's features are appraised for *relevance* to robot goals (relevant or irrelevant) and *valence* to goals (positive or negative outcome expectancies). User features (e.g., intelligence) encoded as positive (e.g., 'helpful') may afford the facilitation of a desired robot goal. This instigates positive outcome expectancy. The comparison between the features of robot and user establishes a level of *similarity* (similar or dissimilar). The measures in the encode phase - mediated by relevance and valence in the comparison phase and moderated by similarity - determine the robot's responses.

In the *response* phase, the robot establishes the levels of *involvement* with and *distance* towards the user. Involvement and distance are two tendencies that occur in parallel and compensate one another [14]. In addition, the robot calculates a value for the so called *use intentions*, the willingness to employ the user again as a tool to achieve robot goals. Together with involvement and distance, the use intentions determine the overall satisfaction of the robot with its user.

Based on this level of *satisfaction*, the robot may decide to continue or stop the interaction and turn to another user. In the Affective Decision Making module [7], the robot makes a decision on the more rationally generated use intentions in unison with the more affectively generated involvement-distance trade-off. The action that promises the highest expected satisfaction during interaction is selected.

3 Triple Comparison

Fig. 1 depicts the similarities and differences between CoMERG, EMA, and I-PEFiC^{ADM}. Although explicitly mentioned in I-PEFiC^{ADM} alone, it is not hard to apply the encode-compare-respond phases to CoMERG and EMA. In the next sections, we offer a comparison of models, using Fig. 1 as our reference point.

3.1. Encode

According to CoMERG, people can select different *situations*, or modify the situation they are currently in, to regulate their emotions. In CoMERG the evaluation of how 'good' a certain situation is, is assumed to be given. In EMA, situations are appraised using the utility of state predicates about the situation and a causal interpretation of this situation. The agents can cope with these situations to change the person-environment-relationship, either by motivating changes to the interpretation of this relationship or by motivating actions that change the environment. I-PEFiC^{ADM} regards the user, agent, or character as part of a situation and that situation primes the features that are selected and how they are perceived.

Features in CoMERG are called aspects. According to CoMERG, a person can focus on one or another aspect (feature) of the world to regulate his or her emotions. In EMA, "current state predicates" in effect relate to those features considered in the subjectively construed situation. State predicates are statements about features of the environment which can be true or false. In I-PEFiC^{ADM}, features receive a certain weight according to frequency of occurrence, salience, or familiarity. Weights can change because of attentional shifts or situation changes.

The *appraisal domains* of I-PEFiC focus on characters. There is a host of evidence that for the judgment of fictional characters [7] and embodied agents [14], users classify features as good or bad, beautiful or ugly, realistic or unrealistic, and as aids or obstacles. According to EMA, on the other hand, agents perceive the world according to a causal interpretation of past and ongoing world events, including plans and intentions of self and others and past actions.

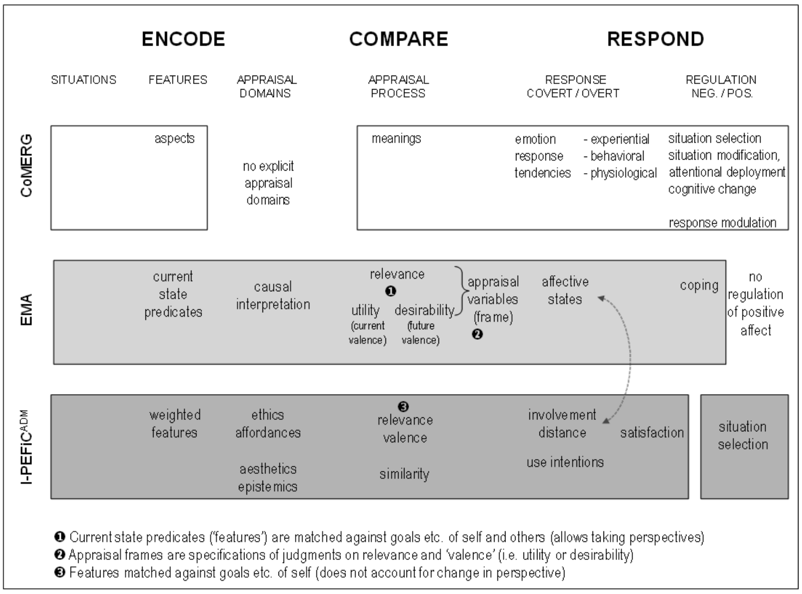


Fig. 1. Overview of CoMERG, EMA, and I-PEFiC^{ADM}

3.2. Compare

CoMERG refers to the *appraisal process* by "cognitive meanings." According to CoMERG, a person can perform the emotion regulation strategy 'cognitive change,' by attaching a different cognitive meaning to a situation. One type of cognitive change, *reappraisal*, means that an individual cognitively re-evaluates a potentially emotion-eliciting situation in terms that decrease its emotional impact [5]. In the view advocated by I-PEFiC^{ADM}, personal meaning is attached to a feature through relevance and valence. In a particular (imagined) situation, an object or feature may potentially benefit or harm someone's goals, beliefs, or concerns and as such, acquires 'meaning' ([4] cf. primary appraisal in [8]).

In EMA, this meaning is acquired through an appraisal process which is modeled in much detail. In this process, multiple appraisal frames are generated to allow for taking different perspectives. These appraisal frames are generated using many appraisal variables, which are taken from the theory of Smith & Lazarus, who call them the appraisal components, and Roseman, who calls them appraisal dimensions. Most of these appraisal variables could be mapped to relevance and valence used in I-PEFiC^{ADM}. According to EMA, relevance measures the significance of an event for the agent. Unlike Frijda, however, EMA equates significance with utility, which in Frijda's terms would be 'valence'. An event outcome is only deemed significant in EMA if it facilitates or inhibits a state predicate with non-zero utility. Valence is not explicitly mentioned in EMA although "utility" and "desirability" can be regarded as two instantiations of it.

Utility is a measure of the relative satisfaction from (or desirability of) environmental features. EMA represents preferences over environmental features as numeric utility over the truth-value of state predicates. Utilities may be either intrinsic (meaning that the agent assigns intrinsic worth to this environmental feature) or extrinsic (meaning that they inherit worth through their probabilistic contribution to an intrinsically valuable state feature). Utility, then, may be viewed as positive or negative outcome expectations about features in the current situation and is expressed in current state predicates (hence, 'current valence').

Desirability covers both a notion of intrinsic pleasantness and goal congruence (in Scherer's typology), as well as a measure of importance or relevance. It captures the appraised valence of an event with regard to an agent's preferences. An event is desirable, from some agent's perspective, if it facilitates a state to which the agent attributes positive utility or if it inhibits a state with negative utility. Like utility, desirability may be viewed as positive or negative outcome expectations but this time about features in the future situation ('future valence').

The explicit division in current and future states is what I-PEFiC^{ADM} is missing, as well as the possibility to change perspectives. EMA and I-PEFiC^{ADM} resemble each other in that causal interpretation of ongoing world events in terms of beliefs, desires, plans, and intentions in EMA is comprised in the beliefs, goals, and concerns that are checked for relevance and valence in I-PEFiC^{ADM}. However, EMA uses a number of variables, called appraisal frames, to cover the appraisal process, whereas in I-PEFiC^{ADM}, these appraisal frames appear to pertain to the more general concepts of relevance and valence. For example, urgency would be a clear-cut specification of relevance (cf. [4]) and ego involvement could be seen as a part of valence. However, EMA also uses some variables (such as causal attribution and coping potential) which are more related to the environment and less to the character, and which are somewhat broader than relevance and valence.

3.3. Respond

Fig. 1 exemplifies that in EMA, relevance of an event as well as utility and desirability (current / future valence) of features are mapped via an appraisal frame onto emotion instances of a particular category and intensity. These are called affective states. This may be seen as a covert response to the situation – an internal affective state that does not yet nor necessarily translate into overt actions. In I-PEFiC^{ADM}, affective states as such are not the focus but rather the involvement-distance trade-off, which is seen as the central process of engagement.

What comes closest to EMA's affective states are involvement and distance (Fig. 1, curved arrows). On this view, emotions emerge during the trade-off. For example, if a girl is asked for a date by a boy she loves, her involvement with him may be accompanied by happiness. When the boy looks at other girls on this date, the girl may still be involved with the boy but this time she feels challenged.

The involvement-distance trade-off could also count as the concretization of the emotion response tendencies that CoMERG hinges on. In CoMERG, these tendencies result in several responses: experiential, behavioral, and physiological. EMA and I-PEFiC^{ADM} are restricted to the experiential and behavioral domain. In EMA, affective states lead to coping behavior. For example, if your car makes strange noises, you might adopt emotion-focused coping (e.g., wishful thinking: tell yourself it is not that important and will probably stop by itself) which will inform the next decision; or you might adopt problem-focused coping to take a specific overt action to address the threat (e.g., have your car checked at the garage). In I-PEFiC^{ADM}, the combination of involvement, distance, and use intentions predicate the level of satisfaction (experiential), which feeds into affective decision making. This results into overt responses (behavior) such as kissing, kicking, or walking away.

CoMERG describes five *emotion regulation strategies* (see Sec. 2.1). Following Gross, CoMERG predicts that strategies that are performed earlier in the process of emotion generation are more effective to regulate one's emotions. EMA provides a more specific model which focuses (in much detail) on coping. Situation selection and situation modification are implemented in EMA via problem-focused coping strategies (i.e., take-action) and avoidance. Attentional deployment corresponds to EMA's strategies of seek/suppress information. Cognitive change corresponds to EMA's various emotion-directed strategies. EMA does not model suppression. I-PEFiC^{ADM} focuses on situation selection. Another difference is that CoMERG and I-PEFiC^{ADM} allow the regulation of affect by increasing, maintaining, or decreasing the positive or negative response, whereas EMA focuses on decreasing negative affect alone. In EMA, being overenthusiastic is left uncontrolled, whereas in CoMERG and I-PEFiC^{ADM}, positive affect can be down-regulated or compensated for. As a result, one can state that coping in EMA is one of the instantiations of emotion regulation in CoMERG.

For EMA, there must be an explicit causal connection between coping strategies and the emotions they are regulating whereas for CoMERG that is not a prerequisite. In CoMERG, people perform strategies to change their level of emotion, which are simply modeled via difference equations. EMA gives a more detailed and formal description of how emotion regulation works. For example, reappraisal as a general emotion regulation strategy in CoMERG is in EMA described in terms of a change in causal interpretation.

4 Integration

In our attempt to integrate the above models, we will adhere to the naming convention of 'features' instead of 'aspects' that the agent can detect in a situation because both EMA and I-PEFiC^{ADM} use that concept and it is interchangeable with 'aspects' in CoMERG. Only I-PEFiC^{ADM} explicitly mentions the appraisal domains that are important in perceiving features. Therefore, the agent will use ethics, affordances, aesthetics, and epistemics as the main domains through which features are funneled into the appraisal process.

CoMERG, EMA, and I-PEFiC^{ADM} all assume or elaborate an appraisal process. CoMERG is least explicit and the concept of 'meaning' can easily be attached to 'personal significance' and 'personal relevance' in both EMA and I-PEFiC^{ADM}. In EMA and I-PEFiC^{ADM}, relevance and valence play an active role, but EMA models the different manifestations rather than the general concepts. In unison, we will use the term relevance to indicate importance or meaning to (dynamic) personal goals, concerns, beliefs, intentions, plans, etc. and valence as (current) utility or (future) desirability of features in a situation. This may instantiate in the form of, for example, urgency as an aspect of relevance and likelihood or unexpectedness as an aspect of valence.

On the response side, EMA focuses on mood and emotions whereas I-PEFiC^{ADM} emphasizes the more general trends of involvement, distance, and use intentions. Yet, they are two sides of the coin that could be called 'affective states.' Emotions and moods may evolve from involvement-distance trade-offs and both the specific (e.g., happy emotions) and general experiential response (e.g., involvement) may be liable to regulation strategies.

CoMERG provides the most profound distinctions with respect to the type of responses (experiential, behavioral, and physiological) and the number of regulation strategies. However, in no way are these distinctions at odds with EMA or I-PEFiC^{ADM}. Coping is best worked out by EMA and situation selection by I-PEFiC^{ADM}, encompassing a module for affective decision making that on the basis of expected satisfaction chooses from several domain actions.

Fig. 2 shows a blueprint for the integration of CoMERG, EMA, and I-PEFiC^{ADM} into a framework for computerized affect generation and regulation. On the far left of the figure, we see a virtual agent. She can perform attentional deployment to weigh the features of her interaction partners. The agent develops state predicates about others in a certain context. Features receive indices for different appraisal domains. The observed other acquires personal meaning or significance for the agent because she compares their features with her personal goals, beliefs, and concerns. This establishes the relevance and valence of others to her goals and concerns. While relevance determines the intensity of affect, valence governs its direction. The agent can also look at others through the eyes of another agent.

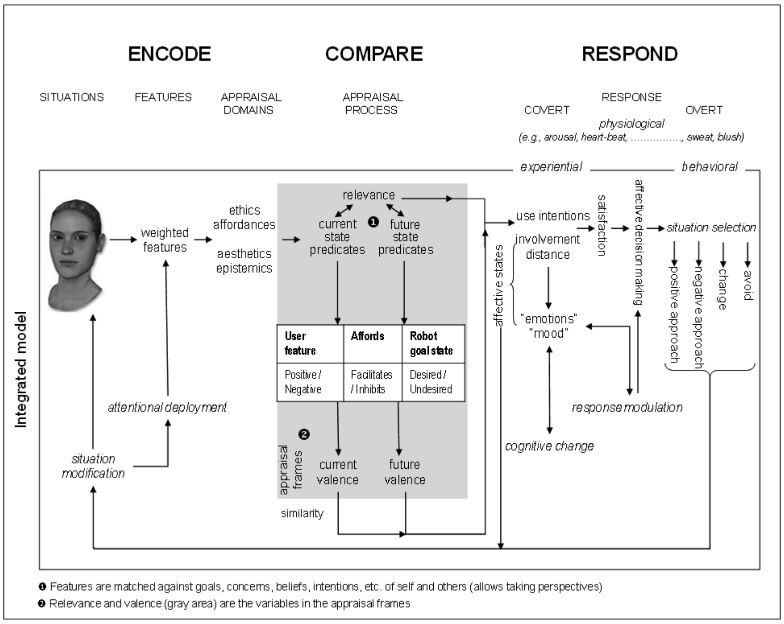


Fig. 2. Proposed integration of the three models

When the (initial) appraisal process is completed, the agent is ready to affectively respond. Relevance, current and future valence form an appraisal frame that feeds into her (un)willingness to 'use' someone for her purposes (e.g., having a conversation) and that helps her to trade friendship (involvement) for keeping her cool (distance). Inside, the agent will 'experience' several (perhaps ambiguous) emotions. On a physiological level, she may be aroused (e.g., increased heart-rate). All this is not visible to others yet; they are covert responses.

During affective decision making, the agent selects the option that promises the highest expected satisfaction. This may be accompanied by physiological reactions such as blushing and trembling. Response modulation may influence the affective decision making. The performed action leads to a new situation.

5 Conclusion

Various researchers from different fields have proposed formal models that describe the processes related to emotion elicitation and regulation (e.g., [2, 3, 7, 10]). For this reason, it is impossible to provide a complete comparison of existing models within one paper. Instead, the approach taken in this article was to select three of the more influential models, which share that they can be used to enhance believability of virtual characters: CoMERG, EMA, and I-PEFiC^{ADM}. The theories by which they were inspired cover most psychological literature in affect-related processes, including the works of Frijda [4], Lazarus [8], and Gross [5].

In this article, we have argued that each of the three approaches has its specific focus. For example, CoMERG covers a wide variety of emotion regulation strategies, whereas I-PEFiC^{ADM} provides an elaborated mechanism for encoding of different appraisal domains, which have empirically shown to be crucial in human-robot interaction. EMA on its turn contains very sophisticated mechanisms for both appraisal and coping, which have already proved their value in various applications. Because several of these features are complementary to each other, this paper explores the possibilities to integrate them into one combined model of affect for virtual humans. For a first attempt to implement this integrated model, see [11]. As a next step, it is planned to perform systematic user-tests in order to assess whether our integration indeed results in more human-like affective behavior than the three sub-models do separately.

References

1. Bosse, T., Pontier, M.A., Siddiqui, G.F., and Treur, J.: Incorporating Emotion Regulation into Virtual Stories. In: Pelachaud, C., Martin, J.C., Andre, E., Chollet, G., Karpouzis, K., and Pele, D. (eds.), Proc. of the Seventh Int. Conference on Intelligent Virtual Agents, IVA'07, pp. 339-347 (2007)
2. Bosse, T., Pontier, M.A., and Treur, J.: A Dynamical System Modeling Approach to Gross´ Model of Emotion Regulation. In: Lewis, R.L., Polk, T.A., and Laird, J.E. (eds.), Proc. of the 8th Int. Cf. on Cognitive Modeling, ICCM'07, pp. 187-192 (2007)
3. Breazeal, C.: Emotion and sociable humanoid robots. In E. Hudlika (Ed.), International Journal of Human Computer Interaction, Vol. 59, pp. 119–155 (2003)
4. Frijda, N. H.: The Emotions. New York: Cambridge University (1986)
5. Gross, J.J.: Emotion Regulation in Adulthood: Timing is Everything. Current directions in psychological science, Vol. 10(6), pp. 214-219 (2001)

6. Hoorn, J.F., Pontier, M.A., and Siddiqui, G.F.: When the user is instrumental to robot goals. First try: Agent uses agent. Proceedings of Web Intelligence and Intelligent Agent Technology 2008 (WI-IAT '08), Vol. 2, pp. 296-301 (2008)
7. Konijn, E.A., and Hoorn, J.F.: Some like it bad. Testing a model for perceiving and experiencing fictional characters. Media Psychology, Vol. 7(2), pp. 107-144 (2005)
8. Lazarus, R.S.: Emotion and Adaptation. New York: Oxford University (1991)
9. Mao, W., and Gratch, J.: Evaluating a computational model of social causality and responsibility. 5th Int. Joint Conference on Autonomous Agents and Multiagent Systems, Hakodate, Japan (2006)
10. Marsella, S., and Gratch, J.: EMA: A Model of Emotional Dynamics. Cognitive Systems Research, Vol. 10(1), pp. 70-90 (2009)
11. Pontier, M.A., and Siddiqui, G.F.: Silicon Coppélia: Integrating Three Affect-Related Models for Establishing Richer Agent Interaction. Proc. of Web Intelligence and Intelligent Agent Technology 2009 (WI-IAT '09), Vol. 2, pp.279-284 (2009)
12. Smith, C.A., and Lazarus, R.S.: Emotion and Adaptation. In L. A. Pervin (Ed.), Handbook of Personality: theory & research, pp. 609-637. NY: Guilford Press (1990)

PART III INTEGRATION OF APPRAISAL, INVOLVEMENT AND REGULATION

CHAPTER 6

Silicon Coppélia: Integrating Three Affect-Related Models for Establishing Richer Agent Interaction

This chapter appeared as Pontier, M., and Siddiqui, G. F., Silicon Coppélia: Integrating Three Affect-Related Models for Establishing Richer Agent Interaction. In: Baeza-Yates, R., Lang, J., Mitra, S., Parsons, S., and Pasi, G. (eds.), Proceedings of the 9th IEEE/WIC/ACM International Conference on Intelligent Agent Technology, IAT'09. IEEE Computer Society Press, 2009, pp. 279-284.

Silicon Coppélia: Integrating three affect-related models for establishing richer agent interaction

Matthijs Pontier and Ghazanfar F. Siddiqui

VU University Amsterdam, Center for Advanced Media Research Amsterdam
Buitenveldertselaan 3, 1082 VA Amsterdam, The Netherlands
{mpr210, ghazanfa}@few.vu.nl

Abstract. Affect modeling plays a vital role in faithfully simulating human emotion and in emotionally-evocative technology. Current affect models are still strong simplifications compared to human affective complexity. To establish richer agent interaction, we integrated three affect-related models: CoMERG, I-PEFiCADM and EMA. These models partly overlap, and where distinct, they complement one another. The integrated model called Silicon Coppélia was implemented and simulation experiments were performed to test the behavior of the model. These experiments show that the model can simulate richer agent behaviors than any of the models could have done alone.

1 Introduction

Compared to human affective complexity, contemporary emotion models of software agents are quite simple. In anticipation of more productive interactions between user and agent, this paper presents an integration of three models that are sufficiently alike, while maintaining their individual qualities. We used three models (CoMERG, EMA, and I-PEFiCADM) of agent affect-generation and affect-regulation that in our view are suitable for integration purposes.

To accomplish less simplistic affective behavior in agents, we took three models inspired by some of the most leading theories in the emotion field. Smith and Lazarus' [8] theory of emotion was formalized by Gratch and Marsella [7] into EMA, a model to create agents that exhibit and cope with (negative) affect. Bosse, Pontier, and Treur [2] used the emotion regulation theory of Gross [5] to develop CoMERG. This model can simulate different emotion regulation strategies explained by Gross. Hoorn, Pontier, and Siddiqui [6] used the concern-driven theory of Frijda [4] to build I-PEFiCADM, a model for building agents that can trade rational for affective choices. These theories are fit for integration due to their adequate mechanisms, simplicity and coherence. Collectively, their foundation (Frijda, Smith & Lazarus, Gross) covers a large part of emotion theory. Because all three were inspired by the appraisal model of emotion, they smoothly fit together. As indicated earlier [1], all three approaches point at important aspects of human affective behavior and all three approaches miss out on something. CoMERG [2] and EMA [7], address the regulation of affective states, but EMA does not regulate positive affect, and it cannot be used to simulate irrational choices where appropriate. CoMERG, on the other hand, has no provisions for generating affect, and does not explicitly account for a causal interpretation of the world state. I-PEFiCADM [6] generates and balances affect but is mute about the different regulation mechanisms.

Because the models are complementary to each other, it makes sense to integrate them. Because EMA, compared to the other models, is more complex and domain-specific, we simplified this model, thereby preserving its core principles. By combining and integrating these models into Silicon Coppélia, this integrated model should be able to simulate richer agent behavior than any of them can do alone. We will test whether this really is the case by performing simulation experiments on Silicon Coppélia under various parameter settings.

2 Implementation

This chapter will describe the implementation of Silicon Coppélia. Figure 1 shows Silicon Coppélia in graphical format. Due to space limitations, not all the formulas in the model will be described, but we will mainly focus on the new formulas in the model compared to the implementation of the affective decision making model I-PEFiCADM [6]. According to I-PEFiCADM, an agent perceives another agent in terms of ethics (good/bad), aesthetics (beautiful/ugly), affordances (aid/obstacle) and realism (cf. [9]).

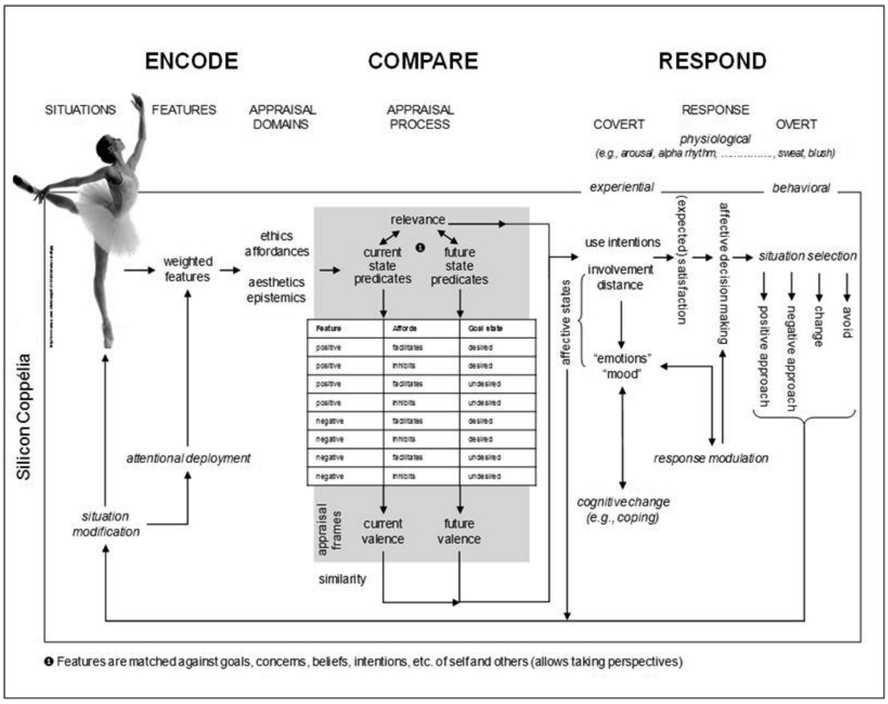


Figure 1: Silicon Coppélia is the integration of CoMERG, EMA, and I-PEFiC^{ADM}

Together with similarity and expected utility this leads via relevance and valence to feelings of involvement and distance, and the more rational use intentions, which are combined into an expected satisfaction for each action. The complete Silicon Coppélia including all the formulas can be found in [10].

2.1. Appraisal frame

Agents are appraised via an appraisal frame. This appraisal frame includes a number of appraisal variables, such as beliefs and expected utilities.

The **believed likelihood** that a goal-state will be accomplished in the range [-1, 1] is calculated by looking at which sub-goals have already been reached. In the system, agents have beliefs about whether states in the world are true in the range [0, 1]. The agents also have beliefs about states facilitating or inhibiting other states in the world in the range [-1, 1] where -1 means an agent believes a state strongly inhibits another state, and 1 means a state strongly facilitates another state. The likelihood of a goal-state is calculated by taking all states into account which are believed to influence the chances of reaching the goal-state. All the beliefs about the states being true are multiplied with the beliefs about the states facilitating or inhibiting the goal-state. The following algorithm is performed to the resulting values. Because other algorithms used in this system have a similar form, we will call it algorithm A:

1. Sort the values in two lists: [0→1] and [0→-1]
2. Start with 0 and take the mean of the value you have and the next value in the list. Continue until the list is finished. Do this for both the negative and the positive list.
3. Outcome = weighed mean of the outcomes of both lists, with weights (#pos / #tot) for the list of positive values, and (#neg / #tot) for the list of negative values.

This way, if, for example, multiple sub-goals that facilitate the goal-state have been accomplished, each sub-goal that is achieved increases the perceived likelihood of reaching the goal-state. However, the more sub-goals have been reached, the less impact each sub-goal has on the perceived likelihood of achieving the goal-state.

In the system, agents have **beliefs that actions facilitate world-states** in the domain [-1, 1] where -1 means an action strongly inhibits a world-state, and 1 means an action strongly facilitates a world-state. If an agent observes an action being performed, and it believes that this action facilitates a certain world-state, it changes its **beliefs that the agent performing the action is responsible for reaching that world-state** according to the following formulas. Because other formulas used in the system have a similar form, we will call them formulas of the form F. Due to space limitations, only one formula of this type is completely shown in the paper. The other formulas of this form can be found in [10].

IF obs(A1, A2, performs, action)
AND belief(action, facilitates, goal-state) > 0
→ belief(A2, responsible, goal-state) = old_belief + mf_{bel_resp} * belief(action, facilitates, goal-state)*(1 – old_belief)

In formulas of the form F, mf<variable> is a modification factor that determines how quickly the variable is updated. This modification factor is multiplied with the impact value, in this case belief(action, facilitates, goal-state). Multiplying with limiter (1-old_belief) manages the formula does not go out of range, and manages that if an agent's belief approaches an extreme value, it will be harder to push it further to the extreme, and easier to get it back to a less extreme value.

Agents have **beliefs about the praiseworthiness** of other agents in the domain [-1, 1] where -1 is very blameworthy and 1 is very praiseworthy. If an agent has a goal and it believes that the goal should already have been reached while it has not, it blames or praises the agents who it believes are responsible for (not) reaching the goal according to formula F with as impact value belief(A1, A2, responsible, goal-state) * ambition_level(goal-state). This way, if the belief about responsibility and ambition level are both positive or negative, A1 will increase its perceived praiseworthiness of A2. If a goal-state is reached, the agent praises or blames the agents that are believed to be responsible for this in a similar manner.

Expected utilities are calculated the same way as in [6]. If an agent has multiple expected utilities of a feature or an action with regard to several goals, a **general expected utility** is calculated by **taking all goals into account** which are believed to be influenced by the feature or action. All the expected utilities of the feature or action with respect to a goal are given to algorithm A, which calculates the general expected utility. The general expected utilities of actions generate **action tendencies** in the agent with the same value.

Compared to [6], now the actions have a continuous level of positivity and negativity in approach, instead of being classified as a certain type. This allows for differentiating between, for example, changing in a positive way (supportive critique) and changing in a negative way (running down on someone). To calculate **general positivity and general negativity in the action tendencies**, all action tendencies are multiplied with the positivity of the action and the negativity of the action in two separate lists. Algorithm A is performed to both these lists to calculate the general positivity and negativity of the action tendencies of the agent. Where in [6] the action tendencies of each class would have a direct effect on involvement and distance, in this paper the general positivity and negativity in action tendencies have an effect on **involvement** and **distance** via **relevance** and valence, as this has been refined in [9]. The use intentions of an agent towards another agent are calculated by performing algorithm A for all features of an agent and actions that can be performed to that agent. The **expected satisfaction** is calculated by trading involvement for distance, and taking a weighed mean of the involvement-distance tradeoff and the use intentions as described in [6]. The agent with the highest expected satisfaction will be picked. Once an agent has been selected, the action to perform to that agent is picked using the following formula:

$$\text{ExpectedSatisfaction}_{(A1,\ Action,\ A2)} = w_{eu} * \text{Action\_Tendency} + w_{pos} * (1 - \text{abs}(\text{positivity} - \text{bias}_I * \text{Involvement})) + w_{neg} * (1 - \text{abs}(\text{negativity} - \text{bias}_D * \text{Distance}))$$

The agent will search for the action with the level of positivity that is closest to the level of (biased) involvement, the level of negativity closest to the level of (biased) distance, and the strongest action tendency. The importance of positivity, negativity and expected utility for selecting an action can be adjusted by changing the weight. Using biases in this process can be seen as a type of **response modulation**.

2.2. Effects on Emotions

Perceived state predicates and appraisal variables lead to emotions. **Hope** and **fear** are based on the believed likelihood that (un)desired world-states will take place. For all world-states with a believed likelihood, the following function is performed to calculate the hope for a goal. This function is similar to the function described in [3].

IF f >= likelihood
→ hope_for_goal = -0,25 * (cos(1/f * π * likelihood(goal)) -1,5) * ambition(goal)

IF f < likelihood
→ hope_for_goal = -0.25 * (cos(1/(1-f)*π*(1-likelihood(goal))))-1.5) * ambition(goal)

Here, f is a shaping parameter (in the domain [0, 1]) that can be used to manipulate the location of the top of the hope curve. The value of this parameter may differ per individual, and represents 'fatalism' (or pessimism): the top of the likelihood/hope-curve is always situated at the point where likelihood = f. In this paper, f is set at 0.5. Algorithm A is performed to the found values for hope_for_goal. Only here, instead of step 3, hope is the outcome of the list with positive values, and fear is the absolute outcome of the list of negative values.

If a world-state becomes true or false, the levels of **joy** and **distress** are calculated by performing formula F with ambition_level(world-state) or a negation thereof as impact value. This way, a desired world-state becoming true will increase joy, and decrease distress, and an undesired does the opposite. This same rule is applied for world-states facilitating other world-states, with belief(state, facilitates, goal-state) * ambition_level(goal-state), or a negation of this multiplication as *impact value*.

If a world-state becomes true, the agent's level of **surprise** will move towards the believed unlikelihood of this world-state happening, using the following formula:

$$\text{surprise} = p_{surp} * \text{old\_surprise} + (1-p_{surp}) * (1\text{–likelihood})$$

In this formula, p_{surp} is a persistency factor, which determines the slowness of adjustment of surprise. If an agent believes that a goal-state should have been reached, but it has not been reached, this will increase its surprise according to formula F with likelihood(goal-state) as *impact value*. Because being surprised does not last forever, it is multiplied with a decay factor each timestep, which is set at 0.95 for all agents.

To calculate the level of **anger** in the range [0, 1] from agents towards other agents, formula F is used with Belief(A2, responsible, goal-state) * Ambition_level(A1, goal-state) as *impact value*. This way, if an agent believes a desired goal-state should have been reached, but it has not, the agent will get more angry at the agents who are believed to be responsible for not reaching this goal-state, and less angry at the agents who are believed to have tried helping reaching the goal-state. If the goal-state was undesired, this is the other way around. Because people do not stay angry forever, it is multiplied with a decay factor each timestep. The general level of anger is calculated by performing algorithm A for all levels of anger from an agent to other agents. Because there is only a list of positive values, the general level of anger simply is the outcome of step 2. **Guilt** is calculated by taking the value of anger at self.

All agents have a desired level of emotion for each type of emotion, similar as in [2]. This desired level will usually be high for positive emotions (joy, hope) and low for negative emotions (anger, guilt). The **overall mood** is calculated by the following formula:

$$\text{Mood} = 1\text{-}(\Sigma(\beta_{emotion} * \text{abs}(\text{Emotion} – \text{desired}(\text{Emotion}))$$

2.3. Emotion Regulation Strategies

To regulate their emotions, agents can perform **situation selection** and **situation modification** by affectively selecting situations and sub-situations with the highest expected satisfaction. **Attentional deployment** can be performed to change the focus of attention. Agents have beliefs that certain features cause emotions. If an agent has attention for a certain feature, and an emotion increases, it will increase its belief that that certain feature causes that particular emotion using formula F with (Emotion(t)–Emotion(t-1)) * Attention(Feature) as *impact value*. In this formula, Emotion(t) is an experienced level of emotion at a certain timepoint (e.g., the level of joy at timepoint 5). Using the belief that a feature Feat causes an emotion E, an agent can shift its attention Att as an emotion regulation strategy using the following formula:

Att(Feat) = old_value – belief(Feat, causes, E)*(E–desired(E))

This formula manages that if an agent believes that a feature causes an emotion, it will increase attention to this feature if it wants to increase its level of this emotion, and decrease its attention if it wants to decrease the level of the emotion. Each timestep, the attention is also shifted based on the relevance of features. The relevance of features is calculated by taking the absolute value of the general expected utility of a feature using the following formula:

Att(Feat) = p_{att} * old_value + (1 - p_{att}) * Relevance(Feat)

In this formula, p_{att} is a persistency factor, which determines the slowness of adjustment of attention. Each timestep, the sum of the levels of attention is normalized to 1.

Cognitive change is implicitly performed by changing beliefs during the simulation as described earlier in this paper. These changes in belief indirectly influence the agents' mood. Cognitive change can also be performed explicitly by changing the causal interpretation of past events, a form of emotion-focused coping. If an agent feels guilty for not reaching a desired goal-state (i.e., the level of guilt is above some threshold), either because the agent performed an action that inhibited it, or it did not perform an action that facilitated it, it can decrease its belief that the action had an influence on reaching the goal by multiplying it with a modification factor.

3 Simulation Experiments

To test the behavior of our model, we implemented the model in JavaScript and performed a number of simulation experiments under different parameter settings. Each experiment concerns a scenario involving a mother, a father and a daughter. There is a party, and the possible actions that were inserted in the system are: 1) going to the party, 2) allowing another agent to go to the party, 3) allowing another agent to go to the party with some restrictions, and 4) forbidding another agent to go to the party. The possible world-states (which the agents can have as goals and/or sub-goals) were 'daughter is having fun', 'daughter is safe', and 'parents are happy'. The results of the experiments are described below. Due to space limitations, not all the performed experiments are described in the paper. The remaining experiments and their simulation traces can be found in [10].

Baseline condition:

To start, an initial experiment was performed that served as a control condition for the remaining experiments. In this condition, all parameters were set to 0, and the biases in perceiving features were set to the neutral value of 1. The complete parameter settings and results for the baseline condition and the experiments can be found in [10].

Experiment 1: Mother gets angry

In this experiment, the mother observes that the father allows the daughter to go to the party. The mother wants her daughter to be safe and believes that allowing the daughter to go to the party strongly inhibits this goal. This belief leads to a negative expected utility for allowing her daughter to go to the party with respect to the goal of having her daughter safe, which leads to a negative action tendency for this action. Compared to the baseline experiment, this decreases her expected satisfaction of allowing her daughter to go to the party. Because the mother observes the father allowing their daughter to go to the party, and she believes that this inhibits the goal of their daughter being safe, she believes the father is responsible for their daughter not being safe. Because she wants her daughter to be safe, she thinks the father is blameworthy, and decreases her view on his ethics. The mother ends up being angry at the father.

Experiment 2: Belief that states lead to other states

In this experiment, the daughter wants her parents to be happy. She thinks that if she is safe and is having fun, this will make her parents feel happy. Due to some external events, at timepoint 1 the daughter is having fun, and at timepoint 2 the daughter is also being safe, which results in the parents being happy at timepoint 3.

At timepoint 1, because she is having fun, the daughter believes that her parents might become happy. Because of this, she has hope for her parents becoming happy, which increases her general level of hope. Also, because the daughter is having fun, and none of the agents had any expectations that this would happen, their level of surprise increases, and the mood of all the agents is increased.

At timepoint 2, because she is having fun and is safe, the daughter believes even stronger that her parents might become happy. Because of this higher likelihood, she is pretty confident that her parents will be happy, and therefore she is not hoping that much anymore as before. Also, because the daughter is safe, and none of the agents had any expectations that this would happen, their level of surprise increases even more, and their mood increases.

At timepoint 3, the parents are even more surprised because they are being happy. The daughter, however, was already expecting her parents to become happy, so her level of surprise decreases. Because her parent being happy was a desired goal of the daughter, her level of joy and her mood also increase.

Experiment 3: Affect overrides rationality

In this experiment, the daughter is a good, beautiful, realistic agent. The father wants the daughter to be safe, and thinks forbidding her to go to the party will facilitate this goal. This leads the father to have a high expected utility for forbidding his daughter to go to the party, and a high action tendency for this action. Due to the changed designed features compared to the baseline condition, there are changes in perceived similarity, relevance and valence. This increases the distance, but especially the involvement of the father towards his daughter. Because of this, the father increases his involvement-distance tradeoff towards his daughter. The expected satisfaction of forbidding his daughter to go to the party increases for the father because of the high expected utility of this action. However, due to the increase in involvement, the expected satisfaction of allowing his daughter to go to the party with restrictions increases even more.

Therefore, the father ends up allowing his daughter to go to the party with restrictions, where rationally he would have chosen to forbid his daughter to go to the party.

4 Discussion

In this paper we presented an implementation of an integration of CoMERG [2], I-PEFiCADM [6], and a simplified version of EMA [7]. Compared to the model in [6], the agents have goal-related beliefs that lead to emotions. Also, some emotion regulation strategies were added to the system based on [3, 9].

Simulation experiments were performed to test the behavior of Silicon Coppélia. In experiment 1, the mother became angry at the father, because in her view, he performed an action (allowing their daughter to go to a party) that was in conflict with her goals (their daughter being safe). In experiment 2, several world-states becoming true led the daughter to having beliefs about the likelihood of other world-states becoming true. This caused her to experience hope, and later in the simulation joy when the expected and desired goal-state became true. Further, all the agents experienced surprise when world-states unexpectedly became true. In experiment 3, due to involvement with his daughter, the father made the affective decision to allow his daughter to go to the party, where rationally he would have forbidden his daughter to go to this party. These results are as would have been expected from the theory [4, 5, 8].

In previous studies, CoMERG and I-PEFiCADM alone [2, 6] were not able to simulate this kind of behavior. I-PEFiCADM and CoMERG are not able to simulate emotions based on beliefs about the responsibility of other agents and the likelihood of goal-states happening, as happens in experiment 1 and 2. EMA [7], on its turn, cannot be used to make irrational decisions where appropriate, as happens in experiment 3. Therefore, we can conclude that the simulation experiments show that Silicon Coppélia can simulate richer agent behavior than CoMERG, I-PEFiCADM or EMA can do alone.

Acknowledgements

We would like to thank Tibor Bosse, Edwin Zwanenburg, Johan F. Hoorn and Jonathan Gratch for their input to this paper.

References

1. Bosse, T., Gratch, J., Hoorn, J.F., Pontier, M.A., and Siddiqui, G.F.: (submitted). “Coppélius' concoction: Similarity and complementarity among three affect-related agent models.” *International Conference on Agents and Artificial Intelligence*, ICAART 2010.
2. Bosse, T., Pontier, M.A., and Treur, J.: “A Dynamical System Modeling Approach to Gross´ Model of Emotion Regulation.” In: Lewis, R.L., Polk, T.A., and Laird, J.E. (eds.), Proc. of the 8th Int. Cf. on Cognitive Modeling, ICCM'07, pp. 187-192 (2007)
3. Bosse, T., and Zwanenburg, E.: “There's Always Hope: Enhancing Agent Believability through Expectation-Based Emotions.” In: Pantic, M., Nijholt, A., Cohn, J. (eds.) *Proceedings of the 2009 International Conference on Affective Computing and Intelligent Interaction, ACII'09. IEEE Computer Society Press, to appear, 2009.*
4. Frijda, N.H.: *The Emotions*, Cambridge University, New York, 1986.
5. Gross, J.J.: “Emotion Regulation in Adulthood: Timing is Everything”, *Current directions in psychological science*, Vol. 10, No. 6, 2001, pp. 214-219.
6. Hoorn, J.F., Pontier, M.A., and Siddiqui, G.F.: “When the user is instrumental to robot goals. First try: Agent uses agent.” *Proceedings of IEEE/WIC/ACM Web Intelligence and Intelligent Agent Technology 2008 (WI-IAT ‘08)*, IEEE/WIC/ACM, Sydney AU, 2008, pp. 296-301.
7. Marsella, S., and Gratch, J.: “EMA: A Model of Emotional Dynamics.” *Cognitive Systems Research*, Vol. 10(1), 2009, pp. 70-90.
8. Smith, C.A., and Lazarus, R.S.: “Emotion and Adaptation.” In: L.A. Pervin (Ed.), *Handbook of Personality: theory & research*, Guilford Press, NY, 1990, pp. 609-637.
9. Van Vugt, H.C., Hoorn, J.F., and Konijn, E.A.: (in press). “Interactive engagement with embodied agents: An empirically validated framework.” *Computer Animation and Virtual Worlds Journal*.
10. See appendix at the end of the chapter (http://www.few.vu.nl/~ghazanfa/iat2009_appendix.html).

Appendix A: Formulas used in the model

Variable	Meaning	Range
abs(X)	The absolute value of variable or formula X	-
max(X, Y)	The maximum value of variables or formulas X and Y	-
$\text{Perceived}_{(<Feature>, A1, A2)}$	New value Agent1 perceives of a certain feature of Agent2	[0, 1]
$\text{Designed}_{(<Feature>, A2)}$	Value assigned by 'the designer' to a certain feature of Agent2	[0, 1]
$\text{Bias}_{(A1, A2, <Feature>)}$	Bias that Agent1 has of Agent2 regarding a certain feature	[0, 2]
Bias_{inv}	bias for desired involvement to show in behavior	[0, 2]
Bias_{dis}	bias for desired Distance to show in behavior	[0, 2]
$\text{Similarity}_{(A1, A2)}$	Similarity of agent towards other agent	[0, 1]
$\text{Dissimilarity}_{(A1, A2)}$	Dissimilarity of agent towards other agent	[0, 1]
$\text{Relevance}_{(A1, A2)}$	Relevance of agent towards other agent	[0, 1]
$\text{Irrelevance}_{(A1, A2)}$	Irrelevance of agent towards other agent	[0, 1]
$\text{Pos\_valence}_{(A1, A2)}$	Positive Valence of agent towards other agent	[0, 1]
$\text{Neg\_valence}_{(A1, A2)}$	Negative Valence of agent towards other agent	[0, 1]
$\text{Involvement}_{(A1, A2)}$	Involvement of agent towards other agent	[0, 1]
$\text{Distance}_{(A1, A2)}$	Distance of agent towards other agent	[0, 1]
$\text{Inv\_Dist\_trade\_off}_{(A1, A2)}$	Outcome of involvement distance trade off	[0, 1]
$\text{Use\_intentions}_{(A1, A2)}$	Use_intentions of agent towards other agent	[0, 1]
γ_{inv_dist}	Weight used in calculating Involvement-Distance Tradeoff	[0, 1]
$\beta_{factor \leftarrow factor}$	regression weight a perceived factor has for another perceived factor of an agent	[-1, 1]
$\text{ExpectedUtility}_{(<Feature>, Agent, Goal)}$	The expected utilities of features	[-1, 1]
$\text{ExpectedUtility}_{(Action, Agent, Goal)}$	The expected utilities of actions	[-1, 1]
$\text{GEU}_{(Feature, Agent)}$	The general expected utility of a feature	
$\text{GEU}_{(Action, Agent}$	The general expected utility of an action	
$\text{Facilitates}_{(<Feature>, Agent, Goal)}$	Belief that a feature of another agent facilitates a goal-state	[-1, 1]
$\text{Facilitates}_{(Action, Agent, Goal)}$	Belief that an action of another agent facilitates a goal-state	[-1, 1]
$\text{agent\_bel\_likelihood}_{(Agent, Statel)}$	Belief about the likelihood that something will happen	[-1, 1]
$\text{agent\_bel\_agent\_responsible}_{(Agent, Other\_Agent, State)}$	Belief that other agent is responsible for reaching a goal	[-1, 1]
$\text{agent\_bel\_agent\_praiseworthy}_{(Agent, Other\_Agent)}$	Belief that other agent is praiseworthy	[-1, 1]
$\text{Action\_Tendency}_{(Agent, Action, Other\_Agent)}$	Agents Action_Tendency for an action towards another agent	[-1, 1]
$\text{Positivity}_{(Action)}$	Level of positivity of an action	[0, 1]
$\text{Negativity}_{(Action)}$	Level of negativity of an action	[0, 1]
$\text{Expected\_Satisfaction}_{(Agent, Other\_agent)}$	Expected_Satisfaction of performing an action towards another agent	[0, 1]
$\text{Emotion}_{(Agent, <Emotion>)}$	Level of emotion of an agent	[0, 1]

Variable	Meaning	Range
$Desired_{<Emotion>}$	Desired emotion of an agent	[0, 1]
$Mood_{(Agent)}$	Level of the mood of an agent	[0, 1]
F	Fatalism used for calculating hope and fear	[0, 1]
$Rel_{(Feature, Agent)}$	Relevance of a feature of an agent	[0, 1]
$Attention_{(Feature, Agent)}$	Attention level for a feature of an agent	[0, 1]
$mf_{<Factor>}$	Modification factor that determines the speed of change when updating a factor	[0, 1]
$p_{<Factor>}$	Persistency factor for updating the value of a factor	[0, 1]
$Decay_{<Factor>}$	Decay factor for updating the value of a factor	[0, 1]
α_{guilt_bc}	Belief update factor when performing cognitive change based on guilt	[-1,1]
$Threshold_{guilt}$	Threshold of guilt for performing cognitive change	[-1,1]

To increase readability, the names of the beta weights use the following acronyms:

Variable	Acronym	Variable	Acronym
Good	good	Similarity	sim
Bad	bad	Dissimilarity	ds
Beautiful	bea	Relevance	rel
Ugly	ugly	Irrelevance	irr
Realistic	real	Positive Valence	pv
Unrealistic	unr	Negative Valence	nv
Aid	aid	Involvement	inv
Obstacle	obst	Distance	dis

Calculating perceived feature values:

All formulas are in the form Perceived(<feature>, A1, A2) = Bias(A1, A2, <feature>) * Designed(<feature>, A2):

Perceived(Beautiful, A1, A2) = Bias(A1, A2, Beautiful) * Designed(Beautiful, A2)
Perceived(Ugly, A1, A2) = Bias(A1, A2, Ugly) * Designed(Ugly, A2)

Perceived(Realistic, A1, A2) = Bias(A1, A2,Realistic) * Designed(Realistic, A2)
Perceived(Unrealistic, A1, A2) = Bias(A1, A2, Unrealistic) * Designed(Unrealistic, A2)

Perceived(Good, A1, A2) = Bias(A1, A2,Good) * Designed(Good, A2)
Perceived(Bad, A1, A2) = Bias(A1, A2, Bad) * Designed(Bad, A2)

Perceived(Intended_aid, A1, A2) = Bias(A1, A2, Intended_Aid) * Designed(Intended_Aid, A2)
Perceived(Intended_obst, A1, A2) = Bias(A1, A2, Intended_Obst) * Designed(Intended_Obst, A2)

Updating bias for perceiving ethics:

IF Belief(A1, goal-state, should_be_reached)= true
AND belief(A1, goal-state, false)
THEN Bias(A1, A2, Good)=old_bias(A1, A2,Good) – mf $_{bias\_ethics}$ * old_bias(A1, A2, Good)
Bias(A1, A2, Bad)=old_bias(A1, A2, Bad) + mf $_{bias\_ethics}$ * (2 - old_bias(A1, A2, Bad))

Calculating expected utilities of features and actions

ExpectedUtility_feature(Feature, Agent, Goal) = Belief(facilitates(Feature, Agent, Goal)) * Ambition(Goal-state)
ExpectedUtility_action(Action, Agent, Goal) = Belief(facilitates(Action, Agent, Goal)) * Ambition(Goal-state)

Calculating a General Expected Utility of a feature of an agent

1. Sort all expected utilities of features in two lists: [0-->1] and [0-->-1]
2. Start with 0 and take the mean of the value you have and the next value in the list. Continue until the list is finished. Do this for both the negative and the positive list.
3. GEU(Feature, Agent) = weighed mean of the outcomes of both lists, with weights (#pos / #tot) for the list of positive values, and (#neg / #tot) for the list of negative values.

Calculating a General Expected Utility of performing action towards an agent

1. Sort all expected utilities of actions in two lists: [0-->1] and [0-->-1]
2. Start with 0 and take the mean of the value you have and the next value in the list. Continue until the list is finished. Do this for both the negative and the positive list.
3. GEU(Action, Agent) = weighed mean of the outcomes of both lists, with weights (#pos / #tot) for the list of positive values, and (#neg / #tot) for the list of negative values.

Calculating Action Tendencies:

Action_Tendency(Action, Agent) = GEU(Action, Agent)

Calculating the General Positivity and Negativity Action Tendencies

1a. Take all Action_Tendency(Action, Agent) * Positivity(Action)
1b. Take all Action_Tendency(Action, Agent) * Negativity(Action)
2. Sort all Action_Tendencies in two lists: [0-->1] and [0-->-1]
3. Start with 0 and take the mean of the value you have and the next value in the list. Continue until the list is finished. Do this for both the negative and the positive list.
4a. GPAT(agent, other_agent) = weighed mean of the outcomes of the positive and the negative list of 1a, with weights (#pos / #tot) for the list of positive values, and (#neg / #tot) for the list of negative values
4b. GNAT(agent, other_agent) = weighed mean of the outcomes of the positive and the negative list of 1b, with weights (#pos / #tot) for the list of positive values, and (#neg / #tot) for the list of negative values

Calculating Similarity, Dissimilarity, Relevance, Irrelevance, Positive Valence and Negative Valence:

$$
\begin{aligned}
\text{Similarity}_{(A1, A2)} = {} & 1 - (\\
& \beta_{sim \leftarrow good} * \text{abs}(\text{Perceived}_{(Good, A1, A2)} - \text{Perceived}_{(Good, A1, A1)}) + \\
& \beta_{sim \leftarrow bad} * \text{abs}(\text{Perceived}_{(Bad, A1, A2)} - \text{Perceived}_{(Bad, A1, A1)}) + \\
& \beta_{sim \leftarrow bea} * \text{abs}(\text{Perceived}_{(Beautiful, A1, A2)} - \text{Perceived}_{(Beautiful, A1, A1)}) + \\
& \beta_{sim \leftarrow ugly} * \text{abs}(\text{Perceived}_{(Ugly, A1, A2)} - \text{Perceived}_{(Ugly, A1, A1)}) + \\
& \beta_{sim \leftarrow real} * \text{abs}(\text{Perceived}_{(Realistic, A1, A2)} - \text{Perceived}_{(Realistic, A1, A1)}) + \\
& \beta_{sim \leftarrow unr} * \text{abs}(\text{Perceived}_{(Unrealistic, A1, A2)} - \text{Perceived}_{(Unrealistic, A1, A1)})
\end{aligned}
$$

$$
\begin{aligned}
\text{Dissimilarity}_{(A1, A2)} = {} & \\
& \beta_{ds \leftarrow good} * \text{abs}(\text{Perceived}_{(Good, A1, A2)} - \text{Perceived}_{(Good, A1, A1)}) + \\
& \beta_{ds \leftarrow bad} * \text{abs}(\text{Perceived}_{(Bad, A1, A2)} - \text{Perceived}_{(Bad, A1, A1)}) + \\
& \beta_{ds \leftarrow bea} * \text{abs}(\text{Perceived}_{(Beautiful, A1, A2)} - \text{Perceived}_{(Beautiful, A1, A1)}) + \\
& \beta_{ds \leftarrow ugly} * \text{abs}(\text{Perceived}_{(Ugly, A1, A2)} - \text{Perceived}_{(Ugly, A1, A1)}) + \\
& \beta_{ds \leftarrow real} * \text{abs}(\text{Perceived}_{(Realistic, A1, A2)} - \text{Perceived}_{(Realistic, A1, A1)}) + \\
& \beta_{ds \leftarrow unr} * \text{abs}(\text{Perceived}_{(Unrealistic, A1, A2)} - \text{Perceived}_{(Unrealistic, A1, A1)})
\end{aligned}
$$

$$
\begin{aligned}
\text{Relevance}_{(A1, A2)} = {} & \\
& \beta_{rel \leftarrow good} * \text{Perceived}_{(Good, A1, A2)} + \\
& \beta_{rel \leftarrow bad} * \text{Perceived}_{(Bad, A1, A2)} + \\
& \beta_{rel <\text{--} pos} * (GAT_{pos}+1)/2 + \\
& \beta_{rel <\text{--} neg} * (GAT_{neg}+1)/2
\end{aligned}
$$

$\text{Irrelevance}_{(A1, A2)} = 1 - ($
$\beta_{irr\leftarrow good} * \text{Perceived}_{(Good, A1, A2)} +$
$\beta_{irr\leftarrow bad} * \text{Perceived}_{(Bad, A1, A2)}) +$
$\beta_{irr<\text{--}pos} * (GAT_{pos}+1)/2 +$
$\beta_{irr<\text{--}neg} * (GAT_{neg}+1)/2$

$\text{Positive\_Valence}_{(A1, A2)} =$
$\beta_{pv\leftarrow good} * \text{Perceived}_{(Good, A1, A2)} +$
$\beta_{pv\leftarrow bad} * \text{Perceived}_{(Bad, A1, A2)} +$
$\beta_{pv<\text{--}pos} * (GAT_{pos}+1)/2 +$
$\beta_{pv<\text{--}neg} * (GAT_{neg}+1)/2$

$\text{Negative\_Valence}_{(A1, A2)} =$
$\beta_{nv\leftarrow good} * \text{Perceived}_{(Good, A1, A2)} +$
$\beta_{nv\leftarrow bad} * \text{Perceived}_{(Bad, A1, A2)} +$
$\beta_{nv<\text{--}pos} * (GAT_{pos}+1)/2 +$
$\beta_{nv<\text{--}neg} * (GAT_{neg}+1)/2$

<u>Calculating Involvement and Distance:</u>

$\text{Involvement}_{(A1, A2)} =$
$\beta_{inv\leftarrow bea} * \text{Perceived}_{(Beautiful, A1, A2)} +$
$\beta_{inv\leftarrow ugly} * \text{Perceived}_{(Ugly, A1, A2)} +$
$\beta_{inv\leftarrow real} * \text{Perceived}_{(Realistic, A1, A2)} +$
$\beta_{inv\leftarrow unr} * \text{Perceived}_{(Unrealistic, A1, A2)} +$
$\beta_{inv\leftarrow aid} * \text{Perceived}_{(Intended\_aid, A1, A2)} +$
$\beta_{inv\leftarrow ob} * \text{Perceived}_{(Intended\_obst, A1, A2)} +$
$\beta_{inv\leftarrow pv} * \text{Pos\_Valence}_{(A1, A2)} +$
$\beta_{inv\leftarrow nv} * \text{Neg\_Valence}_{(A1, A2)} +$
$\beta_{inv\leftarrow ps} * \text{Pos\_Valence}_{(A1, A2)} * \text{Similarity}_{(A1, A2)} +$
$\beta_{inv\leftarrow ns} * \text{Neg\_Valence}_{(A1, A2)} * \text{Similarity}_{(A1, A2)} +$
$\beta_{inv\leftarrow pd} * \text{Pos\_Valence}_{(A1, A2)} * \text{Dissimilarity}_{(A1, A2)} +$
$\beta_{inv\leftarrow nd} * \text{Neg\_Valence}_{(A1, A2)} * \text{Dissimilarity}_{(A1, A2)} +$
$\beta_{inv\leftarrow pb} * \text{Pos\_Valence}_{(A1, A2)} * \text{Perceived}_{(Beautiful, A1, A2)} +$
$\beta_{inv\leftarrow nb} * \text{Neg\_Valence}_{(A1, A2)} * \text{Perceived}_{(Beautiful, A1, A2)} +$
$\beta_{inv\leftarrow pu} * \text{Pos\_Valence}_{(A1, A2)} * \text{Perceived}_{(Ugly, A1, A2)} +$
$\beta_{inv\leftarrow nu} * \text{Neg\_Valence}_{(A1, A2)} * \text{Perceived}_{(Ugly, A1, A2)} +$
$\beta_{inv\leftarrow rel} * \text{Relevance}_{(A1, A2)} +$
$\beta_{inv\leftarrow irr} * \text{Irrelevance}_{(A1, A2)} +$
$\beta_{inv\leftarrow rs} * \text{Relevance}_{(A1, A2)} * \text{Similarity}_{(A1, A2)} +$
$\beta_{inv\leftarrow is} * \text{Irrelevance}_{(A1, A2)} * \text{Similarity}_{(A1, A2)} +$
$\beta_{inv\leftarrow rd} * \text{Relevance}_{(A1, A2)} * \text{Dissimilarity}_{(A1, A2)} +$
$\beta_{inv\leftarrow id} * \text{Irrelevance}_{(A1, A2)} * \text{Dissimilarity}_{(A1, A2)} +$
$\beta_{inv\leftarrow rb} * \text{Relevance}_{(A1, A2)} * \text{Perceived}_{(Beautiful, A1, A2)} +$
$\beta_{inv\leftarrow ib} * \text{Irrelevance}_{(A1, A2)} * \text{Perceived}_{(Beautiful, A1, A2)} +$
$\beta_{inv\leftarrow ru} * \text{Relevance}_{(A1, A2)} * \text{Perceived}_{(Ugly, A1, A2)} +$
$\beta_{inv\leftarrow iu} * \text{Irrelevance}_{(A1, A2)} * \text{Perceived}_{(Ugly, A1, A2)}$

$$\begin{aligned}
\text{Distance}_{(A1,\,A2)} = \; & \beta_{dis \leftarrow bea} * \text{Perceived}_{(Beautiful,\,A1,\,A2)} + \\
& \beta_{dis \leftarrow ugly} * \text{Perceived}_{(Ugly,\,A1,\,A2)} + \\
& \beta_{dis \leftarrow real} * \text{Perceived}_{(Realistic,\,A1,\,A2)} + \\
& \beta_{dis \leftarrow unr} * \text{Perceived}_{(Unrealistic,\,A1,\,A2)} + \\
& \beta_{dis \leftarrow aid} * \text{Perceived}_{(Intended_aid,\,A1,\,A2)} + \\
& \beta_{dis \leftarrow ob} * \text{Perceived}_{(Intended_obst,\,A1,\,A2)} + \\
& \beta_{dis \leftarrow pv} * \text{Pos_Valence}_{(A1,\,A2)} + \\
& \beta_{dis \leftarrow nv} * \text{Neg_Valence}_{(A1,\,A2)} + \\
& \beta_{dis \leftarrow ps} * \text{Pos_Valence}_{(A1,\,A2)} * \text{Similarity}_{(A1,\,A2)} + \\
& \beta_{dis \leftarrow ns} * \text{Neg_Valence}_{(A1,\,A2)} * \text{Similarity}_{(A1,\,A2)} + \\
& \beta_{dis \leftarrow pd} * \text{Pos_Valence}_{(A1,\,A2)} * \text{Dissimilarity}_{(A1,\,A2)} + \\
& \beta_{dis \leftarrow nd} * \text{Neg_Valence}_{(A1,\,A2)} * \text{Dissimilarity}_{(A1,\,A2)} + \\
& \beta_{dis \leftarrow pb} * \text{Pos_Valence}_{(A1,\,A2)} * \text{Perceived}_{(Beautiful,\,A1,\,A2)} + \\
& \beta_{dis \leftarrow nb} * \text{Neg_Valence}_{(A1,\,A2)} * \text{Perceived}_{(Beautiful,\,A1,\,A2)} + \\
& \beta_{dis \leftarrow pu} * \text{Pos_Valence}_{(A1,\,A2)} * \text{Perceived}(_{Ugly,\,A1,\,A2)} + \\
& \beta_{dis \leftarrow nu} * \text{Neg_Valence}_{(A1,\,A2)} * \text{Perceived}(_{Ugly,\,A1,\,A2)} + \\
& \beta_{dis \leftarrow rel} * \text{Relevance}_{(A1,\,A2)} + \\
& \beta_{dis \leftarrow irr} * \text{Irrelevance}_{(A1,\,A2)} + \\
& \beta_{dis \leftarrow rs} * \text{Relevance}_{(A1,\,A2)} * \text{Similarity}_{(A1,\,A2)} + \\
& \beta_{dis \leftarrow is} * \text{Irrelevance}_{(A1,\,A2)} * \text{Similarity}_{(A1,\,A2)} + \\
& \beta_{dis \leftarrow rd} * \text{Relevance}_{(A1,\,A2)} * \text{Dissimilarity}_{(A1,\,A2)} + \\
& \beta_{dis \leftarrow id} * \text{Irrelevance}_{(A1,\,A2)} * \text{Dissimilarity}_{(A1,\,A2)} + \\
& \beta_{dis \leftarrow rb} * \text{Relevance}_{(A1,\,A2)} * \text{Perceived}_{(Beautiful,\,A1,\,A2)} + \\
& \beta_{dis \leftarrow ib} * \text{Irrelevance}_{(A1,\,A2)} * \text{Perceived}_{(Beautiful,\,A1,\,A2)} + \\
& \beta_{dis \leftarrow ru} * \text{Relevance}_{(A1,\,A2)} * \text{Perceived}(_{Ugly,\,A1,\,A2)} + \\
& \beta_{dis \leftarrow iu} * \text{Irrelevance}_{(A1,\,A2)} * \text{Perceived}(_{Ugly,\,A1,\,A2)}
\end{aligned}$$

<u>Calculating Use Intentions</u>

1. Sort all expected utilities about actions in two lists: [0-->1] and [0-->-1]
2. Start with 0 and take the mean of the value you have and the next value in the list. Continue until the list is finished. Do this for both the negative and the positive list.
3. Use_Intentions(Agent, Other_agent) = weighed mean of the outcomes of both lists, with weights (#pos / #tot) for the list of positive values, and (#neg / #tot) for the list of negative values.

<u>Calculating Involvement-Distance Tradeoff</u>

Involvement-Distance-Tradeoff(agent, other_agent) = γ * max(I, D) + (1 - γ) * (I + D) / 2

<u>Calculating Expected Satisfaction</u>

Expected_Satisfaction(Agent,Other_ Agent) = wesidt * Involvement-Distance-Tradeoff(agent, other_agent) +
wesui * Use_Intentions(Agent, Other_agent)

Expected Satisfaction Action =
wesaeu * Action_Tendency +
wesapos * (1 - abs(positivity – bias_inv * Involvement(A1,A2))) +
wesaneg * (1 - abs(negativity – bias_distance * Distance(A1, A2)))

<u>If an agent observes something, it believs it is true</u>

```
IF      agent_obs_state(agent, state) == true)
THEN    agent_beliefs_state(agent, state) = 1;
```

Calculating likelihood of goal-states will be accomplished:

1. Sort the values in two lists: [0-->1] and [0-->-1]
2. Start with 0 and take the mean of the value you have and the next value in the list. Continue until the list is finished. Do this for both the negative and the positive list.
3. likelihood (goal-state) = weighed mean of the outcomes of both lists, with weights (#pos / #tot) for the list of positive values, and (#neg / #tot) for the list of negative values.

Calculating beliefs that agents are responsible for reaching a goal:

IF Obs(A1, A2, performs, action)
AND belief(action, facilitates, goal-state) > 0 (Agent beliefs observed action facilitates goal-state)
THEN belief(A2, responsible, goal-state) = old_belief + mf_{bel_resp} * belief(action, facilitates, goal-state) * (1 – old_belief)

IF Obs(A1, A2, performs, action)
AND belief(action, facilitates, goal-state) < 0 (Agent beliefs observed action will not facilitates goal-state)
THEN Belief(A2, responsible, goal-state) = old_belief + mf_{bel_resp} * belief(action, facilitates, goal-state) * (1 + old_belief)

Calculating beliefs about blameworthy and praiseworthy:

IF World-state changes from false to true
AND ambition (world-state) > 0
THEN belief(A1, A2, praiseworthy, X)=old_belief(A1, A2, praiseworthy, X) + mf_blame * Belief(A2, responsible, goal-state)* ambition (goal-state) * (1- old_belief(A1, A2, praiseworthy, X))

IF World-state changes from false to true
AND ambition (world-state) < 0
THEN belief(A1, A2, praiseworthy, X)=old_belief(A1, A2, praiseworthy, X) + mf_blame * Belief(A2, responsible, goal-state)* ambition (goal-state) * (-1- old_belief(A1, A2, praiseworthy, X))

IF Belief(agent, goal-state, should_be_reached)= true
AND world_state is false
AND ambition (world-state) > 0
THEN belief(A1, A2, praiseworthy, X)=old_belief(A1, A2, praiseworthy, X) - mf_blame * Belief(A2, responsible, goal-state)* ambition (goal-state) * (-1- old_belief(A1, A2, praiseworthy, X))

IF Belief(agent, goal-state, should_be_reached)= true
AND world_state is false
AND ambition (world-state) < 0
THEN belief(A1, A2, praiseworthy, X)=old_belief(A1, A2, praiseworthy, X) - mf_blame * Belief(A2, responsible, goal-state)* ambition (goal-state) * (1- old_belief(A1, A2, praiseworthy, X))

Calculating Emotions:

Calculating Levels for Hope and Fear:

1. Calculate hope_for_goal's using the following rules:
   IF F>=likelihood(goal-state)
   AND (-0.25 * (cos(1 / F * pi * likelihood(goal-state)) -1.5) * ambition(goal-state) >= 0)
   THEN hope_for_goal = -0,25*(cos(1/F * π * likelihood(goal-state)) – 1.5) * ambition(goal-state)

IF F>=likelihood(goal-state)
AND (-0.25 * (cos(1 / F * pi * likelihood (goal-state)) -1.5) * ambition(goal-state) < 0)
THEN hope_for_goal = -0,25*(cos(1/F * π * likelihood(goal-state)) – 1.5) * ambition(goal-state)

IF F<likelihood(goal-state)
AND (-0.25 * (cos(1 / (1-F)* pi * (1-likelihood(goal-state))) -1.5) * ambition(goal-state) >= 0)
THEN hope_for_goal =-0.25*(cos(1/(1-F)*π*(1-likelihood(goal-state)))-1.5)* ambition(goal-state)

IF F<likelihood(goal-state)
AND (-0.25 *(cos(1 / (1-F) *pi*(1-likelihood(goal-state))) -1.5) * ambition(goal-state)< 0)
THEN hope_for_goal =-0.25*(cos(1/(1-F)*π*(1-likelihood(goal-state)))-1.5)*ambition(goal-state)

2. Sort all hope_for_goal's in two lists: [0-->1] and [0-->-1]
3. Start with 0 and take the mean of the value you have and the next value in the list. Continue until the list is finished. Do this for both the negative and the positive list.
4. Outcome positive list = **Hope**, abs(Outcome negative list) = **Fear**

Calculating Joy and Distress:

Based on reaching goals:

IF World-state changes from false to true *(joy and distress based on reaching goal-state)*
AND ambition (world-state) > 0
THEN Joy = old_joy + mf_{joy} * ambition (world-state) * (1 – old_joy)
Distress = old_distress + $mf_{distress}$ * (-ambition (world_state))* old_distress

IF World-state changes from false to true *(joy and distress based on reaching goal-state)*
AND ambition (world-state) < 0
THEN Joy = old_joy + mf_{joy} * ambition (world-state) * old_joy
Distress = old_distress + $mf_{distress}$ *(-ambition (world_state)) * (1 - old_distress)

IF World-state changes from true to false *(joy and distress based on reaching goal-state)*
AND ambition (world-state) > 0
THEN Joy = old_joy + mf_{joy} * -ambition (world_state) * old_joy
Distress = old_distress + $mf_{distress}$ * ambition (world_state) * (1 - old_distress)

IF World-state changes from true to false *(joy and distress based on reaching goal-state)*
AND ambition (world-state) < 0
THEN joy = old_joy + mf_{joy} * -ambition (world_state) * (1-old_joy)
Distress = old_distress + $mf_{distress}$ * ambition (world_state) * old_distress

Based on reaching subgoals:

A goal-state is the same as a world-state. A subgoal is a world-state that is a leads to a goal-state. To increase the readability of the formulas, the subgoal is called world-state, and the goal is called goal-state in the formulas below.

IF World-state changes from false to true
AND ambition (world-state) > 0
THEN Joy = old_joy + mf_{joy} * bel(state, facilitates, goal-state) * ambition (goal-state) * (1-old_joy)
Distress = old_distress + $mf_{distress}$ * -bel(state, facilitates, goal-state) * ambition (goal-state) * old_distress

IF World-state changes from false to true
AND ambition (world-state) < 0
THEN Joy = old_joy + mf_{joy} * bel(state, facilitates, goal-state) * ambition (goal-state) * (1-old_joy)

Distress = old_distress + $mf_{distress}$ * -bel(state, facilitates, goal-state) * ambition (goal-state) * old_distress

IF World-state changes from true to false
AND ambition (world-state) > 0
THEN Joy = old_joy + mf_{joy} * -bel(state, facilitates, goal-state) * ambition (goal-state) * old_joy

Distress = old_distress + $mf_{distress}$ * bel(state, facilitates, goal-state) * ambition (goal-state)* (1 - old_distress)

IF World-state changes from true to false
AND ambition (world-state) < 0
THEN Joy = old_joy + mf_{joy} * -bel(state, facilitates, goal-state) * ambition (goal-state) * (1-old_joy)

Distress = old_distress + $mf_{distress}$ * bel(state, facilitates, goal-state) * ambition (goal-state)* old_distress

Calculating Level for Surprise:

IF World-state changes from false to true
THEN Surprise = $p_{suprise}$ * old_surprise + (1 - $p_{surprise}$) * (1 – likelihood(goal-state))

If a goal-state becomes true, the level of surprise moves towards the believed unlikelihood of this goal-state happening:
IF belief(agent, goal-state, false)
AND Belief(agent, goal-state, should_be_reached)= true
THEN Surprise = old_surprise + $mf_{surprise}$ * likelihood(goal-state) * (1 – old_surprise)

Calculating Levels of Anger and Guilt:

IF Belief(A1, goal-state, should_be_reached) = true
AND world-state is false
THEN Anger_At(A1, A2) =
old_anger_at + mf_{anger} * (- Belief(A2, responsible, goal-state)) * Ambition(A1, goal-state) * (1 - old_anger_at)

Guilt(Agent) = Anger_At(A1, A1)

Calculating general level of anger:

1. Take all levels for anger_at, except for anger_at(A1, A1)
2. Sort all anger_at's in a list
3. Start with 0 and take the mean of the value you have and the next value in the list. Continue until the list is finished.
4. Anger is the outcome of step 3

Aggregate emotions into overall mood:
Mood = 1- (
β_{hope} * abs(Hope – desired(Hope)) +
β_{fear} * abs(Fear – desired(Fear)) +
β_{joy} * abs(Joy – desired(Joy)) +
$\beta_{distress}$ * abs(Distress – desired(Distress)) +
$\beta_{surprise}$* abs(Surprise – desired(Surprise)) +
β_{anger} * abs(Anger – desired(Anger)) +
β_{guilt} * abs(Guilt – desired(Guilt))

Cognitive Change based on Guilt:
If you feel guilty for not reaching a desired goal-state, either because you performed an action that inhibited it, or you did not perform an action that facilitated it, you can decrease your belief that the action had an influence on reaching the goal (this is cognitive change, a form of emotion-focused coping).

If an agent blames itself (in general), and believes he performed an action that inhibits a desired Or facilitates an undesired goal-state, he will decrease the belief that the action influences the goal-state:
IF belief(A1, goal-state, false)
AND belief(A1, A1, praiseworthy, X) < thr_guilt
AND belief(A1, action, facilitates, goal-state) * ambition (goal-state) < 0
AND belief(A1, performed, action) = true
THEN belief(A1, action, facilitates, goal-state) = mf_{guilt} * old_belief (with mf_{guit} at about 0.9)

If an agent blames itself (in general), and believes he did not perform an action that inhibits a desired Or facilitates an undesired goal-state, he will decrease the belief that the action influences the goal-state:
IF belief(A1, goal-state, false)
AND belief(A1, A1, praiseworthy, X) < thr_guilt
AND belief(A1, action, facilitates, goal-state) * ambition (goal-state) > 0
AND belief(A1, performed, action) = false
THEN belief(A1, action, facilitates, goal-state) = mf_{guilt} * old_belief (with mf_{guit} at about 0.9)

Calculating the Relevance of a feature of an agent:

Rel (Feature, Agent) = abs(GEU(Feature, Agent))

Attention changes according to Relevance feature:

Attention(Feature, Agent) = $p_{attention}$ * old_value + (1 - $p_{attention}$) * Rel(Feature, Agent)

Updating beliefs that features cause emotions:

Agents should have believes that certain features cause emotions. Emotion(t) is an experienced level of emotion at a certain timepoint (so for instance, the level of joy at timepoint 5). If an agent has attention for a certain feature, and an emotion increases, it will increase its belief that that feature causes that emotion. mf_{bel_feat} is a modification factor that represents the speed with which beliefs are updated. Dividing with ((1-old_belief)/2) manages that the formula does not make the variable go out of range, and makes it harder to increase your belief if it is already very high.

IF Attention(Feature) * (Emotion(t) – Emotion(t-1)) >=0
THEN Belief(Feature, causes, Emotion) = old_belief + mf_{bel_feat} * (Emotion(t) – Emotion(t-1)) * Attention(Feature) * ((1 – old_belief)/2)

IF Attention(Feature) * (Emotion(t) – Emotion(t-1)) <0
THEN Belief(Feature, causes, Emotion) = old_belief + mf_{bel_feat} * (Emotion(t) – Emotion(t-1)) * Attention(Feature) * ((1 + old_belief)/2)

Attention level for features change according to experiences:

Attention(Feature) = old_value – Belief(Feature, causes, Emotion) * (Emotion – desired_Emotion)

If you belief a feature causes an emotion, you will increase attention to this feature if you want to increase your level of emotion, and decrease attention if you want to decrease the level of emotion.

Sum attention levels of an agent is normalized to 1 each round:

Attention (Feature) = Attention (Feature) / Σ(Attention(Feature))

Appendix B: Parameter settings in the baseline experiment

This appendix contains all parameter settings for the experiments in the paper. These are the parameter settings of the baseline condition. The variables that are changed in other experiments are mentioned in the paper.

Table 1: All values for the regression weights

Weight of X	**on Y**	**Value**
Good	Similarity	0.30
Bad	Similarity	0.20
Beautiful	Similarity	0.20
Ugly	Similarity	0.10
Realistic	Similarity	0.10
Unrealistic	Similarity	0.10
Good	Dissimilarity	0.20
Bad	Dissimilarity	0.30
Beautiful	Dissimilarity	0.10
Ugly	Dissimilarity	0.20
Realistic	Dissimilarity	0.10
Unrealistic	Dissimilarity	0.10
Good	Relevance	0.25
Bad	Relevance	0.20
Positivity	Relevance	0.30
Negativity	Relevance	0.25
Good	Irrelevance	0.20
Bad	Irrelevance	0.25
Positivity	Irrelevance	0.25
Negativity	Irrelevance	0.30
Good	Positive Valence	0.40
Bad	Positive Valence	0.02
Positivity	Positive Valence	0.55
Negativity	Positive Valence	0.03
Good	Negative Valence	0.02
Bad	Negative Valence	0.40
Positivity	Negative Valence	0.03
Negativity	Negative Valence	0.55
Beautiful	Involvement	0.15
Ugly	Involvement	0.05
Realistic	Involvement	0.10
Unrealistic	Involvement	0.05
Aid	Involvement	0.15
Obstacle	Involvement	-0.10
Positive Valence	Involvement	0.55
Negative Valence	Involvement	-0.15
Positive Valence * similarity	Involvement	0.12
Negative Valence * similarity	Involvement	-0.15
Positive Valence * dissimilarity	Involvement	-0.10
Negative Valence * dissimilarity	Involvement	0.05

Positive Valence * beautiful	Involvement	0.07
Negative Valence * beautiful	Involvement	-0.10
Positive Valence * ugly	Involvement	0.07
Negative Valence * ugly	Involvement	-0.04
Relevance	Involvement	0.15
Irrelevance	Involvement	-0.01
Relevance * similarity	Involvement	0.10
Irrelevance * similarity	Involvement	0.04
Relevance * dissimilarity	Involvement	0.03
Irrelevance * dissimilarity	Involvement	-0.02
Relevance * beautiful	Involvement	0.10
Irrelevance * beautiful	Involvement	0.03
Relevance * ugly	Involvement	0.03
Irrelevance * ugly	Involvement	0.01
Beautiful	Distance	-0.05
Ugly	Distance	0.15
Realistic	Distance	0.05
Unrealistic	Distance	0.10
Aid	Distance	-0.10
Obstacle	Distance	0.25
Positive Valence	Distance	-0.35
Negative Valence	Distance	0.40
Positive Valence * similarity	Distance	-0.15
Negative Valence * similarity	Distance	0.20
Positive Valence * dissimilarity	Distance	0.08
Negative Valence * dissimilarity	Distance	-0.05
Positive Valence * beautiful	Distance	0.08
Negative Valence * beautiful	Distance	0.22
Positive Valence * ugly	Distance	-0.05
Negative Valence * ugly	Distance	-0.04
Relevance	Distance	0.15
Irrelevance	Distance	0.05
Relevance * similarity	Distance	-0.08
Irrelevance * similarity	Distance	0.05
Relevance * dissimilarity	Distance	0.05
Irrelevance * dissimilarity	Distance	0.02
Relevance * beautiful	Distance	-0.08
Irrelevance * beautiful	Distance	-0.05
Relevance * ugly	Distance	0.10
Irrelevance * ugly	Distance	0.05
Inv-distance trade-off	Expected satisfaction	0.8
Use_intentions	Expected satisfaction	0.2
Expected Utility	Expected satisfaction action	0.4
Positivity action	Expected satisfaction action	0.3
Negativity action	Expected satisfaction action	0.3
β_{hope}	Mood	0.15
β_{fear}	Mood	0.15
β_{joy}	Mood	0.15
$\beta_{distress}$	Mood	0.15
$\beta_{surprise}$	Mood	0.10
β_{anger}	Mood	0.15
β_{guilt}	Mood	0.15

Table 2: Parameter settings used in the simulation experiments

Variable name	Description of Variable	Value
biasinv	bias for desired involvement to show in behavior	1
biasdis	bias for desired distance to show in behavior	1
mf_respbel	Modification Factor for belief someone is responsible for something	0.25
mf_blame	Modification Factor for belief someone is blameworthy / praiseworthy	0.5
F	Fatalism (or pessimism) for calculating hope and fear	0.5
mf_joy	Modification factor for joy (speed with which joy is updated)	0.25
mf_distress	Modification factor for distress (speed with which distress is updated)	0.25
p_surprise	Persistency for surprise	0.7
mf_surprise	Modification factor of surprise	0.25
decay_surprise	Decay for surprise	0.95
mf_anger	Modification factor for anger	0.5
decay_anger	Decay for anger	0.95
p_guilt	Persistency for guilt	0.25
decay_guilt	Decay for guilt	0.95
thr_guilt	Threshold of guilt for performing cognitive change	-0.8
alpha_guilt_bc	Belief update factor when performing cognitive change based on guilt	0.8
p_att	Persistency for attention to feature	0.9
mf_belief_emotion	Modification factor for belief that feature causes emotion	0.25
mf_bias_ethics	Modification factor of bias on ethics	0.25
γ_{inv_dist}	For involvement distance trade-off	0.5
d_emotion	Desired level of emotion of an agent	0.20
Desired_hope	Desired_emotion	0.80
Desired_fear	Desired_emotion	0.30
Desired_joy	Desired_emotion	0.90
Desired_distress	Desired_emotion	0.10
Desired_surprise	Desired_emotion	0.40
Desired_anger	Desired_emotion	0.10
Desired_guilt	Desired_emotion	0.10

Table 3: Designed values for the features of each agent

Agent	Feature	Value
Mother	Beautiful	0
Mother	Ugly	0
Mother	Good	0
Mother	Bad	0
Mother	Realistic	0
Mother	Unrealistic	0
Father	Beautiful	0
Father	Ugly	0
Father	Good	0
Father	Bad	0
Father	Realistic	0
Father	Unrealistic	0
Daughter	Beautiful	0
Daughter	Ugly	0
Daughter	Good	0
Daughter	Bad	0
Daughter	Realistic	0
Daughter	Unrealistic	0

Table 4: Biases the agents have in perceiving features of their-selves and others

Bias of Agent	**For perceiving**	**Of Agent**	**Value**
Mother	Beautiful	Mother	1
Mother	Beautiful	Father	1
Mother	Beautiful	Daughter	1
Mother	Ugly	Mother	1
Mother	Ugly	Father	1
Mother	Ugly	Daughter	1
Mother	Realistic	Mother	1
Mother	Realistic	Father	1
Mother	Realistic	Daughter	1
Mother	Unrealistic	Mother	1
Mother	Unrealistic	Father	1
Mother	Unrealistic	Daughter	1
Mother	Good	Mother	1
Mother	Good	Father	1
Mother	Good	Daughter	1
Mother	Bad	Mother	1
Mother	Bad	Father	1
Mother	Bad	Daughter	1
Mother	Intended_Aid	Mother	1
Mother	Intended_Aid	Father	1
Mother	Intended_Aid	Daughter	1
Mother	Intended_Obstacle	Mother	1
Mother	Intended_Obstace	Father	1
Mother	Intended_Obstacle	Daughter	1
Father	Beautiful	Mother	1
Father	Beautiful	Father	1
Father	Beautiful	Daughter	1
Father	Ugly	Mother	1
Father	Ugly	Father	1
Father	Ugly	Daughter	1
Father	Realistic	Mother	1
Father	Realistic	Father	1
Father	Realistic	Daughter	1
Father	Unrealistic	Mother	1
Father	Unrealistic	Father	1
Father	Unrealistic	Daughter	1
Father	Good	Mother	1
Father	Good	Father	1
Father	Good	Daughter	1
Father	Bad	Mother	1
Father	Bad	Father	1
Father	Bad	Daughter	1
Father	Intended_Aid	Mother	1
Father	Intended_Aid	Father	1
Father	Intended_Aid	Daughter	1
Father	Intended_Obstacle	Mother	1
Father	Intended_Obstace	Father	1
Father	Intended_Obstacle	Daughter	1
Daughter	Beautiful	Mother	1
Daughter	Beautiful	Father	1
Daughter	Beautiful	Daughter	1
Daughter	Ugly	Mother	1
Daughter	Ugly	Father	1

Daughter	Ugly	Daughter	1
Daughter	Realistic	Mother	1
Daughter	Realistic	Father	1
Daughter	Realistic	Daughter	1
Daughter	Unrealistic	Mother	1
Daughter	Unrealistic	Father	1
Daughter	Unrealistic	Daughter	1
Daughter	Good	Mother	1
Daughter	Good	Father	1
Daughter	Good	Daughter	1
Daughter	Bad	Mother	1
Daughter	Bad	Father	1
Daughter	Bad	Daughter	1
Daughter	Intended_Aid	Mother	1
Daughter	Intended_Aid	Father	1
Daughter	Intended_Aid	Daughter	1
Daughter	Intended_Obstacle	Mother	1
Daughter	Intended_Obstace	Father	1
Daughter	Intended_Obstacle	Daughter	1

Appendix C: Experiments

Baseline Experiment:

To start, an initial experiment was performed that served as a control condition for the remaining experiments. In this condition, all parameters were set to 0, and the biases in perceiving features were set to the neutral value of 1. The desired levels of emotion were set to 0.8 for hope, 0.3 for fear, 0.9 for joy, 0.1 for distress, 0.4 for surprise, 0.1 for anger and 0.1 for guilt for all agents. The positivity and negativity of actions were defined as can be seen in Table 1 for all agents:

Table 1. Positivity and negativity of actions

Action	Positivity	Negativity
Allow to go to party	0.8	0.2
Allow to go with restrictions	0.6	0.4
Forbid to go to party	0.2	0.8
Go to party	0.9	0.1

The complete parameter settings for the baseline condition can be found in [12]. This led to all emotions, perceived feature values, beliefs, expected utilities, action tendencies and general positivity and negativity in the action tendencies being 0. Because all the agents were exactly the same, the perceived similarity was 1 and dissimilarity was 0 for all agents. For all agents, the perceived relevance was 0.28, irrelevance was 0.73, and positive valence as well as negative valence was 0.29. The perceived involvement was 0.20, and the perceived distance was 0.12, leading to an involvement-distance trade-off of 0.18. All use intentions were 0, together with the involvement-distance trade-off leading to an expected satisfaction of 0.24 for all agents. All expected satisfactions for going to the party with another agent were 0.38, while all the expected satisfactions for the other actions the agents could perform were 0.40. The resulting mood level for all the agents was 0.62.

Experiment-1: Mother gets angry

Parameter settings:

Statement	Meaning
agent_ambition(Mother, Daughter_is_safe) = 1	Mother wants her daughter to be safe
belief(Mother, Allow_to_go_to_party, Daughter, facilitates, Daughter_is_safe) = -1	Mother beliefs allowing her daughter to go to the party will strongly inhibit the goal of her daughter being safe
agent_obs_action(Mother, Father, Allow_to_go_to_party, Daughter) = 1	Mother observes the father allowing the daughter to go to the party at timestep 1
agent_bel_should_be_reached(Mother, Daughter_is_safe) = true	Mother beliefs her daughter should be safe starting from timestep 1

Results:

In this experiment, the mother observes that the father allows the daughter to go to the party. The mother wants her daughter to be safe (ambition level set to 1) and believes that allowing the daughter to go to the party strongly inhibits this goal (belief set to -1).

This belief leads to a negative expected utility for allowing her daughter to go to the party with respect to the goal of having her daughter safe, which leads to a negative action tendency of -0.5 for this action. This leads to the mother having negative use intentions towards her daughter with a level of -0.25.

Compared to the baseline experiment, this decreases her expected satisfaction of performing an action towards her daughter from 0.20 to 0.18, and the expected satisfaction of allowing her daughter to go to the party decreases from 0.50 to 0.30.

Because the mother observes the father allowing their daughter to go to the party, and she believes that this inhibits the goal of their daughter being safe, she believes the father is responsible for their daughter not being safe with a level of -0.25.

Because she wants her daughter to be safe, she thinks the father is blameworthy with a level of -0.125, and increases her bias of perceiving the badness of the father to 1.25, and decreases the bias of perceiving the goodness of the father to 0.75. (However, because the designed values for goodness and badness were set to 0 for all agents, the perceived goodness and badness do not change in this experiment).
She also gets angry at the father with a level of 0.238. Because of this, her general anger level is also increased, to a level of 0.119.

Experiment-2: Belief that states lead to other states

Parameter settings:

Statement	Meaning
agent_ambition(Daughter, Parents_are_happy) = 1	The daughter wants her parents to be happy
agent_bel_s_facil_goal(Daughter, Daughter_has_fun, Parents_are_happy) = 1	The daughter believes that if she has fun, this will make her parents happy
agent_bel_s_facil_goal(Daughter, Daughter_is_safe, Parents_are_happy) = 1	The daughter believes that if she is safe, this will make her parents happy

Starting from timepoint 1, the daughter has fun
Starting from timepoint 2, the daughter is safe
Starting from timepoint 3, the parents are happy

Results:

In this experiment, the daughter wants her parents to be happy (ambition level set to 1). She thinks that if she is safe and is having fun, this will make her parents feel happy (both states set to sub-goals of the state parents happy with value 1).

Due to some external events, at timepoint 1 the daughter is having fun, and at timepoint 2 the daughter is also being safe, which results in the parents being happy at timepoint 3.

At timepoint 1, because she is having fun, the daughter believes that her parents might become happy with a likelihood of 0.5. Because of this, she has hope for her parents becoming happy with a value of 0.63, which leads to a general level of hope of 0.31. Also, because the daughter is having fun, and none of the agents had any expectations that this would happen, their level of surprise increase to 0.29. This leads to an increase of mood to 0.64 for the parents, and to 0.69 for the daughter.

At timepoint 2, because she is having fun *and* is safe, the daughter believes that her parents might become happy with a likelihood of 0.75. Because of this higher likelihood, she is pretty confident that her parents will be happy, and therefore she is not hoping that much anymore as before. Therefore, her hope for her parents becoming happy decreases to 0.38, and her general level of hope decreases to 0.19.

Also, because the daughter is safe, and none of the agents had any expectations that this would happen, their level of surprise increases to 0.48. This leads to an increase of mood to 0.65 for the parents. Because the hope of the daughter has decreased, this slightly decreases her mood to 0.68.

At timepoint 3, the parents are even more surprised with a level of 0.60 because they are being happy. The daughter, however, was already expecting their parents would become happy, so her level of surprise decreases to 0.39. Because her parents being happy was a desired goal of the daughter, her level of joy increases to 0.25. This increases her mood to a level of 0.72.

Experiment-3: Affect overrides rationality

Parameter settings:

Statement	Meaning
designed(Daughter, Good) = 1	The daughter is designed to be ethically good
designed(Daughter, Beautiful) = 1	The daughter is designed to be beautiful
designed(Daughter, Realistic) = 1	The daughter is designed to be realistic
agent_ambition(Father, Daughter_is_safe) = 1	The father wants his daughter to be safe
belief(Father, Forbid_to_go_to_party, Daughter, facilitates, Daughter_is_safe) = 1	The father believes forbidding his daughter to go to the party will facilitate the goal of his daughter being safe

Results:

In this experiment, the daughter is a good, beautiful, realistic agent (features set to 1). The father wants the daughter to be safe (ambition level set to 1), and thinks forbidding her to go to the party will facilitate this goal (belief set to 1).

This leads the father to have an expected utility of 1 for forbidding his daughter to go to the party with respect to the goal of his daughter being safe, and an action tendency of 0.5 for this action. Because of this, the general positivity and negativity in his action tendencies are respectively 0.025 and 0.100. The father also has use intentions of 0.25 towards his daughter.

Due to the changed designed features compared to the baseline condition, the perceived similarity and dissimilarity between the parents and the daughter are now both 0.40. The relevance and irrelevance of the daughter for the father change to respectively 0.54 and 0.51. Similarly, the positive valence and negative valence of the father towards the daughter change to respectively 0.55 and 0.52.

This leads the daughter to have a decreased involvement of 0.16 towards her parents. The involvement and distance of the father towards the daughter change to 0.62 and 0.23 respectively. Because of this, the involvement-distance tradeoff of the daughter towards her parents decreases to 0.15, while the father increases his involvement-distance tradeoff towards his daughter to 0.53.

The expected satisfaction of the father for performing an action towards his daughter increases to 0.55. While the expected satisfaction for forbidding her daughter to go to the party decreases to 0.30 for the mother, the expected satisfaction of performing this action increases to 0.50 for the father because of the high expected utility of this action. However, due to the increase in involvement, the expected satisfaction of allowing his daughter to go to the party with restrictions increases even to 0.54. Therefore, the father ends up allowing his daughter to go to the party with restrictions, where rationally he would have chosen to forbid his daughter to go to the party, because that action had an expected utility of 1 for the father, while allowing the daughter to go to the party with restrictions had an expected utility of 0 for the father.

Experiment-4: Parents disagree

Parameter settings:

Statement	Meaning
designed(Daughter, Beautiful) = 1	The daughter is designed to be beautiful
agent_ambition(Mother, Daughter_has_fun) = 1	The mother wants her daughter to have fun
agent_ambition(Mother, Daughter_is_safe) = 1	The mother wants her daughter to be safe
agent_ambition(Father, Daughter_has_fun) = 1	The father wants her daughter to have fun
agent_ambition(Father, Daughter_is_safe) = 1	The father wants her daughter to be safe
belief(Mother, Beautiful, Daughter, facilitates, Daughter_is_safe) = -1	The mother beliefs the beauty of her daughter inhibits the goal of her daughter being safe
belief(Father, Beautiful, Daughter, facilitates, Daughter_has_fun) = 1	The father beliefs the beauty of his daughter facilitates the goal of his daughter having fun
belief(Mother, Allow_to_go_to_party, Daughter, facilitates, Daughter_is_safe) = -1	The mother believes allowing her daughter to go to the party will inhibit the goal of her daughter being safe
belief(Mother, Forbid_to_go_to_party, Daughter, facilitates, Daughter_is_safe) = 1	The mother believes forbidding her daughter to go to the party will facilitate the goal of his daughter being safe
belief(Father, Allow_to_go_to_party, Daughter, facilitates, Daughter_has_fun) = 1	The father believes allowing his daughter to go to the party will facilitate the goal of his daughter having fun
belief(Father, Forbid_to_go_to_party, Daughter, facilitates, Daughter_has_fun) = -1	The father believes forbidding his daughter to go to the party will inhibit the goal of his daughter having fun

agent_bel_should_be_reached(Father, Daughter_has_fun) = true	Father beliefs his daughter should be having fun starting from timestep 2

Results:

In this experiment, the daughter is designed to be a beautiful agent (designed value set to 1). Compared to the baseline experiment, this leads both the parents to increase their attention to the beauty of their daughter, with a value of 0.22, and therefore slightly decrease their attention to other features in the world.

The parents want their daughter to have fun and to be safe (ambition level set to 1). The mother beliefs that her daughter being beautiful inhibits the goal of her daughter being safe (value set to -1), while the father beliefs this facilitates the goal of the daughter having fun (value set to 1). This leads the mother to perceive an expected utility of -1 for the beauty of her daughter regarding her daughter being safe, and a general expected utility of -0.5 for her beauty. The father, however, perceives an expected utility of 1 for the beauty of his daughter regarding the goal of his daughter having fun, and a general expected utility of 0.5 for her beauty.

Further, the mother beliefs that allowing her daughter to go to the party inhibits the goal of her daughter being safe (value set to -1), and forbidding her to go to the party facilitates her daughter being safe. The father beliefs, however, that allowing his daughter to go to the party facilitates the goal of the daughter having fun (value set to 1), and forbidding her to go to the party inhibits the goal of the daughter having fun (value set to -1).

This leads the mother to have an expected utility of -1 for allowing the daughter to go to the party regarding her safety, and generates an action tendency of -0.5 for this action. She has an expected utility of 1 for forbidding her daughter to go to the party regarding her safety, and generates an action tendency of 0.5 for this action. For the father, however, this leads to an expected utility of 1 for allowing his daughter to go to the party regarding her having fun, and generates an action tendency of 0.5 for this action. It also generates an expected utility of -1 for forbidding his daughter to go to the party regarding his daughter having fu, and he generates an action tendency of -0.5 for this action. This leads both the parents to have a general positivity in action tendencies of 0.125, and a general negativity in action tendencies of -0.125.

Because the parents perceive their daughter as beautiful and themselves not, compared to the baseline experiment, their perceived similarity with the daughter decreases to 0.80, and the perceived dissimilarity increases to 0.10. This is also increases their perceived positive valence of their daughter to 0.32, and decreases their perceived negative valence of their daughter to 0.26. This increases their involvement with their daughter to 0.41, and decreases their distance to 0.06. This leads both the parents to have an involvement-distance tradeoff of 0.33 towards their daughter. The mother has intentions to 'use' her daughter of -0.17 and the father has a value of 0.17 for these use intentions. This leads the mother to have an expected satisfaction of 0.34 for interacting with her daughter, and for the father this expected satisfaction of 0.38.

Compared to the baseline experiment, the mother decreases her expected satisfaction of allowing the daughter to go to the party to 0.24, and increases her expected satisfaction of forbidding the daughter to go to the party to 0.51. The father increases his expected satisfaction of allowing the daughter to go to the party to 0.64, and decreases his expected satisfaction of forbidding the daughter to go to the party to 0.11. The expected satisfactions of the other two possible actions to perform to their daughter inserted in the system increase to 0.44 for both the parents.

This leads the mother to forbid her daughter to go to the party, and the father to allow his daughter to go to the party. Therefore, the mother beliefs the father is responsible for inhibiting their daughters' safety (value = -0.25), while she holds herself responsible for facilitating this goal (value = 0.25). The father, on his turn, holds the mother responsible for inhibiting the goal of their daughter having fun (value = -0.25), while he holds himself responsible for facilitating this goal (value = 0.25).

Because the parents both keep performing the same action, the next timestep these values are increased to 0.44 and -0.44. Further, because the father beliefs the daughter should be having fun starting from this timepoint, while this is not the case, he blames the mother for keeping their daughter from having fun with a value of -0.11, while he praises himself for trying to make his daughter have fun with a value of 0.11. Because of this, he gets a bit angry at the mother, with a value of 0.10, and his general level of anger increases to 0.05.

Further, because he focuses relatively much of his attention to the beauty of his daughter while increasing his level of anger, the father increases the belief that beauty causes anger to 0.002, while this belief is only increased to 0.0006 for the other features.

Experiment-5: Father performs cognitive change

Parameter settings:

Statement	**Meaning**
designed(Daughter, Beautiful) = 1	The daughter is designed to be beautiful
agent_ambition(Mother, Daughter_has_fun) = 1	The mother wants her daughter to have fun
agent_ambition(Mother, Daughter_is_safe) = 1	The mother wants her daughter to be safe
agent_ambition(Father, Daughter_has_fun) = 1	The father wants her daughter to have fun
agent_ambition(Father, Daughter_is_safe) = 1	The father wants her daughter to be safe
belief(Mother, Beautiful, Daughter, facilitates, Daughter_is_safe) = -1	The mother beliefs the beauty of her daughter inhibits the goal of her daughter being safe
belief(Father, Beautiful, Daughter, facilitates, Daughter_has_fun) = 1	The father beliefs the beauty of his daughter facilitates the goal of his daughter having fun
belief(Mother, Allow_to_go_to_party, Daughter, facilitates, Daughter_is_safe) = -1	The mother believes allowing her daughter to go to the party will inhibit the goal of her daughter being safe
belief(Mother, Forbid_to_go_to_party, Daughter, facilitates, Daughter_is_safe) = 1	The mother believes forbidding her daughter to go to the party will facilitate the goal of his daughter being safe
belief(Father, Allow_to_go_to_party, Daughter,	The father believes allowing his

facilitates, Daughter_has_fun) = 1	daughter to go to the party will facilitate the goal of his daughter having fun
belief(Father, Allow_to_go_to_party_with_restrictions, Daughter, facilitates, Daughter_has_fun) = 1 belief(Father, Allow_to_go_to_party_with_restrictions, Daughter, facilitates, Daughter_is_safe) = 1	The father believes allowing his daughter to go to the party with some restrictions will facilitate the goal of his daughter having fun as well as his daughter being safe.
belief(Father, Forbid_to_go_to_party, Daughter, facilitates, Daughter_has_fun) = -1	The father believes forbidding his daughter to go to the party will inhibit the goal of his daughter having fun
threshold_guilt = -0.1	The threshold for blameworthiness of self to perform cognitive change is set to -0.1
agent_bel_should_be_reached(Mother, Daughter_has_fun) = true	Mother beliefs her daughter should be having fun at timestep 2
agent_bel_should_be_reached(Father, Daughter_has_fun) = true	Father beliefs his daughter should be having fun starting from timestep 2

Results:

This experiment uses the same parameter settings as experiment 4. Only now, the father beliefs that allowing their daughter to go to the party facilitates all goals inserted in the system. This leads the father to have an expected utility of 1 for allowing the daughter to go to the party with some restrictions regarding both the goals of the daughter having fun *and* being safe. Therefore, he generates an action tendency of 0.75 for this action. This increases his general positivity in action tendencies to 0.25. This increases his perceived relevance towards the daughter to 0.30, and decreases his perceived irrelevance towards her to 0.71. Further, it increases his perceived positive valence of his daughter to 0.36.

This increases his involvement with his daughter to 0.44, and decreases his distance towards her to 0.05. This causes an increase of involvement distance trade-off to the value of 0.34. His intentions to 'use' his daughter are increased to 0.36. This leads to an expected satisfaction of 0.41 for interacting with his daughter.

Compared to experiment 4, the father increases his expected satisfaction of allowing his daughter to go to the party to 0.75, and performs this action. This causes him to belief he is responsible for facilitating all goals inserted in the system (value = 0.25). Also, the mother does not see him anymore as inhibiting their daughter having fun, because she does not have any beliefs about allowing the daughter to go to the party with some restrictions.

Because he keeps performing the same action, the next timestep these values are increased to 0.44. Further, because the father beliefs the daughter should be having fun starting from this timepoint, while this is not the case, and he did not allow her to go to the party without restrictions while he also believed this would facilitate the daughter having fun, as an emotion regulation strategy he decreases the belief that allowing his daughter to go to the party without restrictions would have led to the daughter having fun to 0.8.

PART IV MODELING INVOLVEMENT IN ECONOMICAL CONTEXT

CHAPTER 7

Modeling Greed of Agents in Economical Context

This chapter appeared as Bosse, T., Siddiqui, G.F., and Treur, J., Modelling Greed of Agents in Economical Context. In: García-Pedrajas, N., Herrera, F., Fyfe, C., Benítez, J.M., and Ali, M. (eds.), Proceedings of the 23rd International Conference on Industrial, Engineering & Other Applications of Applied Intelligent Systems, IEA/AIE'10, Part II. Lecture Notes in Artificial Intelligence, vol. 6097. Springer Verlag, 2010, pp. 407-417.

Modeling Greed of Agents in Economical Context

Tibor Bosse[1], Ghazanfar F. Siddiqui[1,2], and Jan Treur[1]

[1] Vrije Universiteit Amsterdam, Department of Artificial Intelligence, De Boelelaan 1081a, 1081 HV Amsterdam, The Netherlands
[2] Quaid-i-Azam University Islamabad, Department of Computer Science, 45320, Pakistan
{tbosse, ghazanfa, treur}@few.vu.nl ghazanfar@qau.edu.pk
http://www.few.vu.nl/~{tbosse, ghazanfa, treur}

Abstract. A classical debate in economics addresses the advantages and drawbacks of modeling from a macroeconomics perspective as opposed to modeling from a microeconomics perspective. Form the latter psychological aspects at an individual level can be taken into account in a differentiated manner. Within computer science and AI, a similar debate exists about the differences between agent-based and population-based modeling. This paper aligns both debates by exploring the differences and commonalities between population-based and agent-based modeling in economical context. A case study is performed on the interplay between individual greed as a psychological concept and global economical concepts. It is shown that under certain conditions agent-based and population-based simulations show similar results.

Keywords: economics, greed, agent-based and population-based modeling.

1 Introduction

Traditionally, macroeconomics addresses the behavior of a world-wide, national or regional economy as a whole [3], whereas microeconomics investigates the economic behavior and decision making of individual agents, for example, consumers, households or firms [11]. Since the latter aims to understand why and how agents make certain economic decisions, various social, cognitive, and emotional factors of human behavior are studied. This has resulted in the emergence of the field of behavioral economics [14]. Although this may be very useful when one wants to analyze the behavior of individual agents, there is some debate about the extent to which it is useful to incorporate these aspects when studying global processes in economics, e.g., [5]. Do personal factors such as risk avoidance, greed, and personal circumstances provide more insight in the global patterns, or can they simply be ignored or treated in a more abstract, aggregated manner? This paper provides some answers to these questions from a computational perspective.

In recent years, various authors have studied processes in economics by building computational models of them, and analyzing the dynamics of these models using agent-based simulation techniques [15]. Ironically, also in the area of agent-based modeling, a debate exists about the pros and cons of two perspectives, namely agent-based and population-based modeling. Agent-based models are often assumed to produce more detailed, faithful behavior, whereas population-based models abstract from such details to focus on global patterns (e.g., [2], [7], and [9]).

Given these similarities between the debate between macro- and microeconomics on the one hand, and the debate between population-based and agent-based modeling on the other hand, it makes sense to align the two debates. Hence, the goal of the current paper is to explore the differences and commonalities between population-based and agent-based modeling in an economical context. This will be done via a case study on the interplay between individual greed and the global economy.

This paper is structured as follows. In Section 2, the existing debate between agent-based and population-based modeling is briefly explained. In Section 3, both an agent-based and a population-based model are introduced for the example domain. In Section 4, a number of simulation results of both models are shown, and the similarities are discussed. Next, Section 5 provides a mathematical analysis on the models. Section 6 concludes the paper with a discussion.

2 Agent-Based versus Population-Based Modeling

The classical approaches to simulation of processes in which larger groups of agents are involved are population-based: a number of groups are distinguished (populations) and each of these populations is represented by a numerical variable indicating their number or density (within a given area) at a certain time point. The simulation model takes the form of a system of difference or differential equations expressing temporal relationships for the dynamics of these variables. Well-known classical examples of such population-based models address ecological processes, for example, predator-prey dynamics (e.g., [6], [12], [13] and [16]), and the dynamics of epidemics (e.g., [1], [6], and [8]). Such models can be studied by simulation and by using analysis techniques from mathematics and dynamical systems theory.

From the more recently developed agent system area it is often taken as a presupposition that simulations based on individual agents are a more natural or faithful way of modeling, and thus will provide better results (e.g., [2] and [7]). Although for larger numbers of agents such agent-based approaches are more expensive computationally than population-based approaches, such a presupposition may provide a justification of preferring their use over population-based approaches, in spite of the computational disadvantages. In other words, they are justified because the results are expected to deviate from the results of population-based simulation, and are considered more realistic. However, in contrast there is another silent assumption sometimes made, namely that for larger numbers of agents (in the limit), agent-based simulations approximate population-based simulations. This would indicate that for larger numbers of agents agent-based simulation just can be replaced by population-based simulation, which would weaken the justification for agent-based simulation discussed above. In, e.g., ([4; 9]), these considerations are explored for the domains of epidemics and crime displacement, respectively. The results put forward in these papers reveal several commonalities between both types of simulation, but also some differences. For example, for some specific parameter settings (concerning population size and rationality of the individual agents, among others), the results of population-based simulation seem to approximate those of agent-based simulation, whereas for other situations some differences can be observed. Furthermore, as could be expected, the computation time of the populations-based simulations is shown to be much lower than that of the agent-based simulation.

In the next sections, similar issues are explored, but this time for a domain within economics. Comparative simulation experiments have been conducted based on different simulation models, both agent-based and population-based.

3 The Agent-Based and Population-Based Simulation Model

In this section, the two simulation models are introduced. First, an agent-based perspective is taken. The main idea behind this model is that the state of the global (world) economy influences the level of greed of the individual agents in the population, which is supposed to relate to the risk level of their investment decisions: in case the economic situation is positive, then people are tempted to take more risk. Moreover, the investment decisions of the individual agents in turn influence the global economy: in case agents become too greedy [10], this is assumed to have a negative impact on the economic situation, for example, due to higher numbers of bankruptcy. In addition, the state of the economy is assumed to be influenced by technological development which is driven by innovation. Inspired by these ideas, the interplay between agents' greed and the global economy is modeled as a dynamical system, in a way that has some similarity to predator-prey models in two variations: agent-based, where each agent has its own greed level, and population-based, where only an average greed level of the whole population is considered.

The agent-based model assumes n heterogeneous agents, which all interact within a certain economy. For each agent k, the individual greed is represented using a variable y_k, and the global economic situation is represented using a variable x. The complete set of variables and parameters used in the model is shown in Table 1.

Table 1. Variables and parameters used in the agent-based model

Variables	x	World economy
	$y^{(1)}, \ldots, y^{(n)}$	Greed of individual agents
	z	Average greed of the agents (i.e., arithmetic mean of all $y^{(k)}$)
	TD	Technological development level
Parameters	a	Growth rate of the economy
	b	Decrease rate of the economy due to average greed
	$c_1, \ldots, c_n$	Growth rate of an agent's greed based on the economy
	$e_1, \ldots, e_n$	Decrease rate of an agent's greed
	inn	Innovation rate

Based on these concepts, a system of difference equations was designed that consists of $n+3$ formulae; here (2) specifies a collection of n equations for each of the n agents, where each agent has its individual values for $y^{(k)}$, c_k and e_k:

(1) Updating the world economy

$$x_{new} = x_{old} + (a * x_{old} - b * x_{old} * z_{old}) * \Delta t$$

(2) Updating the greed of the agents

$$y^{(k)}_{new} = y^{(k)}_{old} + (c_k * b * x_{old} * y^{(k)}_{old} * (2 - y^{(k)}_{old}) / TD_{old} - e_k * y^{(k)}_{old}) * \Delta t \quad \text{(for all agents } k\text{)}$$

(3) Updating the technological development

$$TD_{new} = TD_{old} + inn * TD_{old} * \Delta t$$

(4) Aggregating greed

$$z_{old} = (\Sigma_k \; y^{(k)}_{old}) / n$$

Table 2. Variables and parameters used in the population-based model

Variables	*x*	World economy
	y	Average greed of the population
	TD	Technological development level
Parameters	*a*	Growth rate of the economy
	b	Decrease rate of the economy due to population greed
	c	Growth rate of the population greed based on the economy
	e	Decrease rate of the population greed
	inn	Innovation rate

The population-based dynamical model is similar to the agent-based model, but the difference is that it abstracts from the differences of the individual agents. This is done by replacing the average greed z over all $y^{(k)}$ in formula (1) by one single variable y indicating the greed of the population as a whole, and using a single formula (2), which is only applied at the population level, in contrast to the collection of formulae (2) in the agent-based model, which are applied for all agents separately. The resulting population-based model is shown in Table 2 and in the formulae below.

(1) Updating world economy

$$x_{new} = x_{old} + (a*x_{old} - b*x_{old}* y_{old}) * \Delta t$$

(2) Updating the greed of the population

$$y_{new} = y_{old} + (c*b* x_{old}*y_{old} *(2-y_{old}) / TD_{old} - e*y_{old}) * \Delta t$$

(3) Updating the technological development

$$TD_{new} = TD_{old} + inn* TD_{old} *\Delta t$$

Note that in differential equation format the agent-based and population-based dynamical model can be expressed by n+2, respectively 3 differential equations as shown in Table 3. Moreover, as the innovation rate inn is assumed constant over time, for both cases the differential equation for TD can be solved analytically with solution $TD(t) = TD(0)\ e^{inn\ t}$.

Table 3. The two models expressed by n+2, respectively 3 differential equations

Agent-based model	**Population-based model**
$dx/dt = ax - bxz$ $d\,y^{(k)} /dt = (c_k b\ xy^{(k)}(2- y^{(k)}) / TD) - e_k\, y^{(k)}$ $dTD/dt = inn\ TD$ $z = (\Sigma_k\ y^{(k)})/n$	$dx/dt = ax - bxy$ $dy/dt =(cb\ xy(2-y) / TD) - ey$ $dTD/dt = inn\ TD$

4 Simulation Results

Based on the model introduced above, a number of simulation experiments have been performed under different parameter settings (with population size varying from 2 to 400 agents), both for the agent-based and for the population-based case. Below, a number of them are described. First an agent-based simulation experiment is described. In this first experiment, 25 agents were involved. The initial settings used for the variables and parameters involved in the experiment are shown in Table 4.

Table 4. Initial settings for variables and parameters

Parameter	Value	Variable	Initial value
a	*1.5*	*x*	*5*
b	*5.8*	*y*	random in [*0.2, 0.3*]
c	random in [*0.0260, 0.0274*]	*TD*	*1*
e	random in [*0.85, 0.89*]		
inn	*0.01*	Δt	*0.1*

The results of the simulations are shown in Figure 1a and 1b. In Figure 1a, time is on the horizontal axis and the value of the world economy is represented on the vertical axis. It is evident from the graph that the economy grows as time increases (but fluctuating continuously). Figure 1b shows the individual greed values of all 25 agents. As can be seen they fluctuate within a bandwidth of about 25% with lowest points between about *0.1* and *0.15,* and highest points around *0.45*. The pattern of the average greed over all 25 agents is shown in Figure 1c.

For the population-based simulation, all the parameter settings are the same as in Table 4, except parameters *y*, *c* and *e*. The values for parameters *y*, *c* and *e* used in the population-based simulation were determined on the basis of the settings for the agent-based simulations by taking the average *y*, *c* and *e* for all fifty agents:

$y = (\Sigma_k y_k)/n$ $\qquad c = (\Sigma_k c_k)/n$ $\qquad e = (\Sigma_k e_k)/n$

The results of the population-based simulations are shown in Figure 2a (economy) and 2b (greed). As can be seen from these figures, the results approximate the results for the agent-based simulation. The difference of the world economy for the population-based and agent-based simulation (averaged over all time points) turns out to be 0.112, and the difference between the average greed of the 25 agents in the agent-based simulation and the greed for the population-based simulation is 0.005.

In addition, a number of simulation runs have been performed for other population sizes. Figure 3a displays the (maximum and average) difference between the world economy in the agent-based model and the world economy in the population-based model for various population sizes. Similarly, Figure 3b displays the difference between the average greed in the agent-based model and the greed in the population-based model for various population sizes. The red line indicates the maximum value and the blue line the average value over all time points. As the figures indicate, all differences approximate a value that is close to 0 as the population size increases. Although the results of these particular simulation experiments should not be over-generalized, this is a first indication that for higher numbers of agents, the results of the agent-based model can be approximated by those of the population -based model.

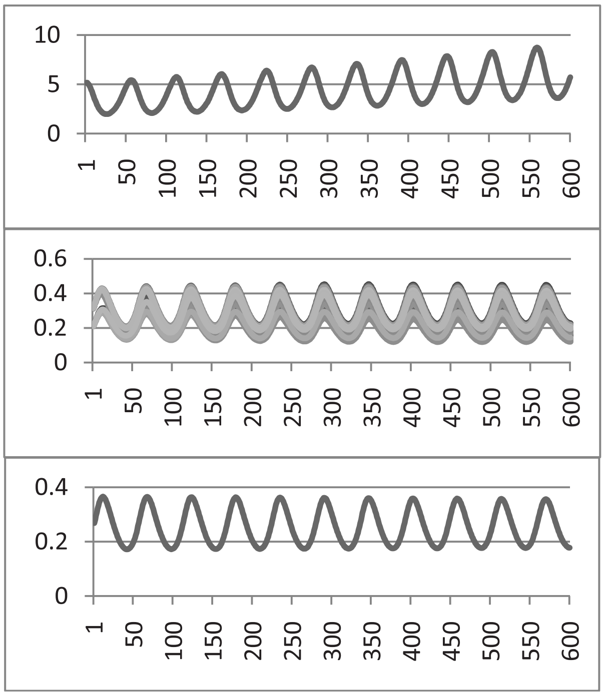


Fig. 1. Agent-based simulation results:
a) world economy, b) individual greed of 25 agents, and c) average greed (over 25 agents)

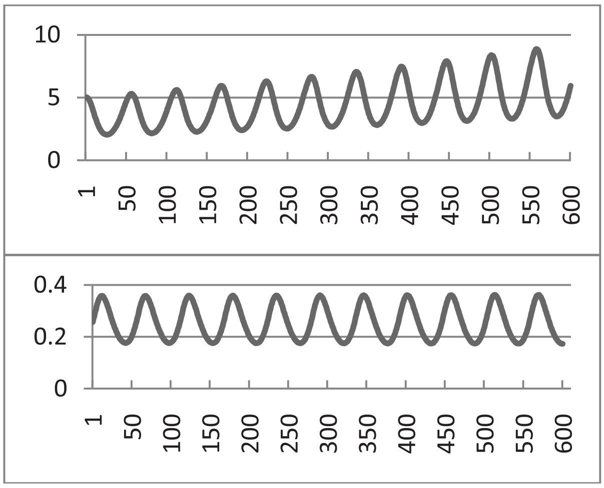


Fig. 2. Population-based simulation results: a) world economy, and b) greed

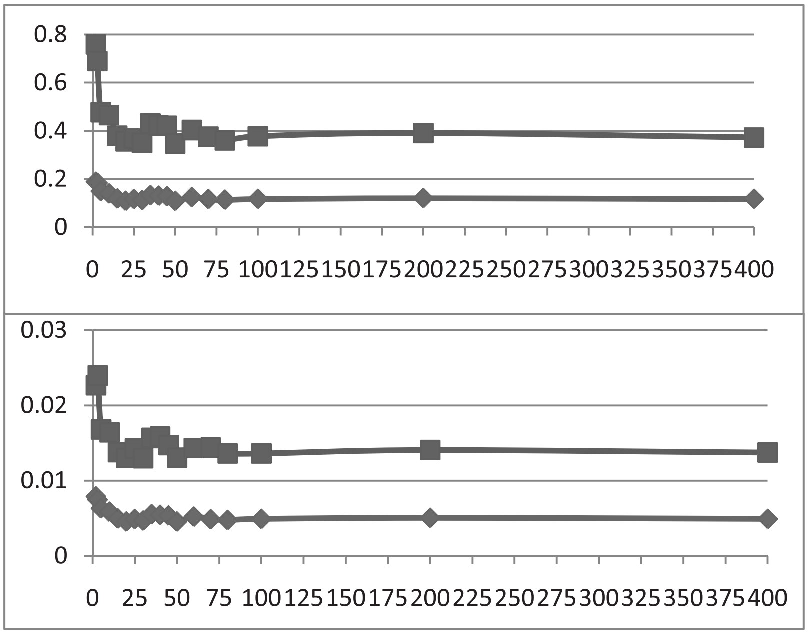


Fig. 3. Difference between both models for various population sizes: a) world economy, and b) greed

5 Mathematical Analysis

In this section a mathematical analysis is presented concerning the conditions under which partial or full equilibria occur; it is assumed that the parameters *a, b, c* and *e* are nonzero. For an overview of the equilibria results, see Table 5.

Dynamics of the economy The economy grows when $dx/dt > 0$ and shrinks when $dx/dt < 0$; it is in equilibrium when $dx/dt = 0$. Assuming x nonzero, according to equation (1) for the population-based model, this can be related to the value of the greed as follows

economy grows $dx/dt > 0 \Leftrightarrow ax - bxy > 0 \Leftrightarrow a - by > 0 \Leftrightarrow y < a/b$
economy shrinks $dx/dt < 0 \Leftrightarrow ax - bxy < 0 \Leftrightarrow a - by < 0 \Leftrightarrow y > a/b$
economy in equilibrium $dx/dt = 0 \Leftrightarrow ax - bxy = 0 \Leftrightarrow a - by = 0 \Leftrightarrow y = a/b$

So, as soon as the greed exceeds a/b the economy will shrink (for example, due to too many bankruptcies), until the greed has gone below this value. This indeed can be observed in the simulation traces. For the agent-based model similar criteria can be derived, but then relating to the average greed z instead of y.

Full Equilibria for the Population-Based Model The first issue to be analyzed is whether (nonzero) equilibria exist for the whole population-based model, and if so, under which conditions. This can be analyzed by considering that x, y and TD are constant and nonzero. For x constant above it was derived from (1) that the criterion is $y = a/b$. For TD constant the criterion is $inn = 0$ as immediately follows from (3). The criterion for $dy/dt = 0$ can be derived from (2) as follows

$$dy/dt = (cbxy(2-y) / TD - ey) = 0 \Rightarrow cbx (2-y) / TD = e \Rightarrow x = (e / ((2b-a) c)) TD$$

This provides the conditions for a full equilibrium

$(1)\quad y = a/b \qquad (2)\ x = (e/((2b\text{-}a)\, c))\ TD \qquad (3)\ inn = 0$

It turns out that for any nonzero setting for the parameters a, b, c and e and for setting $inn = 0$ for the innovation parameter and for any value of TD a nontrivial equilibrium is (only) possible with values as indicated above. Note that this shows that for inn nonzero a nontrivial full equilibrium is not possible, as TD will change over time. However, partial equilibria for greed still may be possible. This will be analyzed next

Equilibria for greed in the population-based model Suppose that the innovation inn is nonzero. In this case it cannot be expected that technological development TD and economy x stay at constant nonzero values. However still for the greed variable y an equilibrium may exist. From the second equation (2) by putting $dy/dt = 0$ it follows

$$cbx\ (2\text{-}y)\ /\ TD = e \ \Rightarrow x = \alpha\ TD \qquad \text{with } \alpha = e\ /\ cb\ (2\text{-}y)$$

By filling this in differential equation (1) it follows

$$d\ \alpha\ TD\ /dt = a\alpha\ TD - b\alpha\ TD\ y \Rightarrow d\ TD\ /dt = (a - by)\ TD$$

By differential equation (3) it can be derived

$$d\ TD\ /dt = (a - by)\ TD = inn\ TD \ \Rightarrow\ (a - by) = inn \ \Rightarrow\ y = (a - inn)/b$$

Note that for $inn = 0$ this also includes the result for the full equilibrium obtained earlier. Moreover, as the equation for TD can be solved analytically, and $x = \alpha\ TD$, also an explicit solution for x can be obtained:

$$TD(t) = TD(0)\ e^{inn\ t} \qquad x(t) = \alpha\ TD(t) = \alpha\ TD(0)\ e^{inn\ t} = x(0)\ e^{inn\ t}$$

Here α can be expressed in the parameters as follows:

$$\alpha = e\ /\ cb\ (2\text{-}y) = e\ /\ cb\ (2\text{-}(a - inn)/b) = (e\ /\ c)\ /\ (2b - a + inn)$$

This shows that according to the model greed can be in an equilibrium $y = (a - inn)/b$, in which case the economy shows a monotonic exponential growth.

Full Equilibria for the agent-based model Similar to the approach followed above:

(1) $dx/dt = (ax - bxz) = 0$
(2) $d\ y^{(k)}\ /dt = (c_k bx\ y^{(k)}\ (2\text{-}\ y^{(k)})\ /\ TD - e_k\ y^{(k)}) = 0$ (for all agents k)
(3) $dTD/dt = inn\ TD = 0$
(4) $z = (\ \Sigma_k\ y^{(k)}\ )/n$

A full equilibrium can be expressed by the following equilibria equations:

(1) $ax = bxz$ (2) $c_k bx\ y^{(k)}\ (2\text{-}\ y^{(k)})\ /\ TD = e_k\ y^{(k)}$
(3) $inn\ TD = 0$ (4) $z = (\Sigma_k\ y^{(k)}\ )/n$

It is assumed that a, b, c_k and e_k are nonzero. One trivial solution is $x = y^{(k)} = 0$. Assuming that x, $y^{(k)}$ and TD all are nonzero, the equations (1) to (3) are simplified:

(1) $a = bz$ (2) $c_k bx\ (2\text{-}\ y^{(k)})\ /\ TD = e_k$ (3) $inn = 0$ (4) $z = (\ \Sigma_k{}^{y(k)}\ )/n$

Th$_i$s provides

(1) $z = a/b$ (2) $y^{(\,)}_k = {}^{2\text{-}}\ e_k\ TD/(c_k bx)$ (3) $inn = 0$ (4) $z = (\ \Sigma_k\ y^{(k)}\ )/n$

From the second, first and last equation it follows that

$$a/b = (\Sigma_k\ y^{(k)}\ )/n = (\Sigma_k\ (2\text{-}\ e_k\ TD/(c_k bx)\ )\ )/n = 2 - \Sigma_k\ (e_k\ TD/(c_k bx)\ )\ /n =$$
$$2 - (TD/bx)\ (\Sigma_k\ (e_k\ /c_k)\ )/n \ \Rightarrow\ x = TD\ \Sigma_k\ (e_k\ /c_k)\ /\ (2b - a)n$$

From this the values for the $y^{(j)}$ can be determined:

$$y^{(j)} = 2\text{-}\ e_j\ TD/(c_j bx) = 2\text{-}\ e_j\ TD/(c_j b\ TD\ \Sigma_k\ (e_k\ /c_k)\ /\ (2b - a)n)$$
$$= 2\text{-}\ e_j\ /(c_j b\ \Sigma_k\ (e_k\ /c_k)\ /\ (2b - a)n) = 2\text{-}\ e_j(2b - a)n\ /(c_j b\ \Sigma_k\ (e_k\ /c_k)\ )$$
$$= 2\text{-}\ e_j(2 - (a/b))n\ /(c_j\ \Sigma_k\ (e_k\ /c_k)\ ) = 2\text{-}\ (2 - (a/b))n\ /(\ \Sigma_k\ (e_k\ /\ e_j\ )(c_j\ /\ c_k)\ )$$

It turns out that for any nonzero setting for the parameters a, b, c_k and e_k and for setting $inn = 0$ for the innovation parameter, and for any value of TD a nontrivial equilibrium is (only) possible with values as indicated above.

Equilibria for greed for the agent-based model From the second equation $c_k bx\ (2-y^{(k)}) / TD = e_k$ with $y^{(k)}$ constant it follows that $x = \alpha_k\, TD$ with α_k the constant $\alpha_k = e_k / c_k\, b\ (2-y^{(k)})$ which aparantly does not depend on k, as both x and TD do not depend on k, so the subscript in α_k can be left out. Filling this in (1) provides:

$$d\,\alpha\, TD/dt = (a\alpha\, TD - b\alpha\, TD\, z) \;\Rightarrow\; d\,TD/dt = (a - bz)\, TD$$

By differential equation (3) it can be derived

$$dTD/dt = (a - bz)\, TD = inn\, TD \;\Rightarrow\; (a - bz) = inn \;\Rightarrow\; z = (a - inn)/b$$

Now the equilibrium values for $y^{(j)}$ can be determined as follows.

$$\alpha = e_k / c_k\, b\ (2-y^{(k)}) \;\Rightarrow\; 2-y^{(k)} = e_k / \alpha\, c_k\, b \;\Rightarrow\; y^{(k)} = 2- e_k / c_k\, \alpha b$$

Next the value of α is determined $z = (\Sigma_k\, y^{(k)})/n = \Sigma_k\, (2- e_k / c_k\, \alpha b)/n = 2- (1/\alpha bn)\, \Sigma_k\, e_k / c_k$. Since $z = (a - inn)/b$ it follows

$$(a - inn)/b = 2- (1/\alpha bn)\, \Sigma_k\, e_k / c_k \;\Rightarrow\; (1/n\alpha)\, \Sigma_k\, e_k / c_k = 2b- (a - inn) \;\Rightarrow$$

$$\Sigma_k\, e_k / c_k = (2b- (a - inn))\, n\alpha \;\Rightarrow\; \alpha = \Sigma_k\, (e_k / c_k) / (2b- (a - inn))n$$

Given this value for α the equilibrium values for the greed $y^{(j)}$ are

$$y^{(j)} = 2- e_j / c_j\, \alpha b = 2- e_j / b\, c_j\, \Sigma_k\, (e_k / c_k) / (2b- (a - inn))n$$

$$= 2- (2- (a - inn)/b)\ \ n / \Sigma_k\, (e_k\, c_j / e_j\, c_k)$$

Table 5. Overview of the equilibria of the two models

	Agent-based model	Population-based model
Full equilibrium	$inn = 0$ $x = (1/(2b - a))\ (\Sigma_k\ (e_k / c_k) / n)\ TD$ $z = a/b$ $y^{(j)} = 2- (2 - (a/b))\ n\ /\ \Sigma_k\ (e_k / e_j)(c_j / c_k)$	$inn = 0$ $x = (1/(2b-a))(\ e/c))\ TD$ $y = a/b$
Partial equilibrium for greed	$TD(t) = TD(0)\ e^{inn\ t}$ $x(t) = (1/(2b- a + inn))\ (\ \Sigma_k\, (e_k / c_k)/n)\ \ TD(0)\ e^{inn\ t}$ $z = (a - inn)/b$ $y^{(j)} = 2- (2- ((a - inn)/b))\ n\ /\ \Sigma_k\, (e_k / e_j)(c_j / c_k)$	$TD(t) = TD(0)\ e^{inn\ t}$ $x(t) = (1/(2b - a + inn))\ (e / c)\ TD(0)\ e^{inn\ t}$ $y = (a - inn)/b$

6 Discussion

This paper discusses similarities and dissimilarities between agent-based models and population-based models in behavioral economics. Inspired by variants of predator prey models (e.g., [6], [12], [13], and [16]), a dynamic behavioral economical model was developed for the relationship between individual agents' greed and the global economy. Simulation experiments for different population sizes were performed for both an agent-based and a population-based model. For both cases the results show that the world economy grows in a fluctuating manner over time and the average greed of the agents fluctuates between *0.1* and *0.45*. A mathematical analysis was performed for both, showing the conditions under which equilibria occur.

It turned out that, in particular for large population sizes, the differences in the economy and average greed between agent-based and population based simulations are close to zero. In different domains, in [4] and [9], under certain conditions similar results were obtained. In literature on agent-based simulation such as in (e.g., [2] and [7]), it is argued that although agent-based modeling approaches are more expensive computationally than population-based modeling approaches, they are preferable due to more accuracy. In contrast to this, the results in the current paper indicate that for the considered domain the agent-based approaches can be closely approximated by population-based simulations. On the other hand, for cases with a rather small number n

of agents the population-based approach may be inadequate. This may raise the question whether a more differentiated point of view in the debate can be considered, namely that for numbers of n agents exceeding a certain N, population-based models are as adequate as agent-based models, whereas for $n < N$ agent-based models are more adequate. A challenge may be to determine this number N for different cases.

For future work, more differentiated personality aspects will be included in the agent model, concerning risk profile and emotions (e.g., feeling insecure) involved, depending upon which decisions are made for the investment (in banking products or stock market). A further aim is to develop a web-based business application incorporating a virtual agent that will interact with a client and regulate the emotions.

References

1. Anderson, R.A., May, R.M.: Infectious Diseases of Humans: Dynamics and Control. Oxford University Press, Oxford, UK (1992)
2. Antunes, L., Paolucci, M.., Norling, E.: Multi-Agent-Based Simulation VIII, Proceedings of the Eighth International Workshop on Multi-Agent-Based Simulation, MABS'07. LNAI, vol. 5003, Springer Verlag (2008)
3. Blanchard, O., Fischer, S.: Lectures in Macroeconomics. MIT Press (1989)
4. Bosse, T., Gerritsen, C., Hoogendoorn, M.., Jaffry, S.W., Treur, J.: Comparison of Agent-Based and Population-Based Simulations of Displacement of Crime. In: Jain, L., et al. (eds.), Proceedings of the 8th IEEE/WIC/ACM International Conference on Intelligent Agent Technology, IAT'08. IEEE Computer Society Press, pp. 469--476 (2008)
5. Brandstätter, H.: Should Economic Psychology Care about Personality Structure? Journal of Economic Psychology, vol.14, pp. 473--494 (1993)
6. Burghes, D.N., Borrie, M.S.: Modeling with Differential Equations. John Wiley (1981)
7. David, N., Sichman, J.S.: Multi-Agent-Based Simulation IX, Proc. of the 9th Int. Workshop on Multi-Agent-Based Simulation, MABS'08. LNAI, vol. 5269, Springer Verlag (2009)
8. Ellner, S.P., and Guckenheimer, J.: Dynamic Models in Biology. Princeton University Press (2006)
9. Jaffry, S.W., Treur, J.: Agent-Based and Population-Based Simulation: A Comparative Case Study for Epidemics. In: L.S. Louca, Y. Chrysanthou, Z. Oplatkova, K. Al-Begain (eds.), Proc. of the 22th European Conference on Modeling and Simulation, ECMS'08. European Council on Modeling and Simulation, pp. 123--130 (2008)
10. Kasser, T., Cohn, S., Kanner, A., Ryan, R.: Some costs of American corporate capitalism: A psychological exploration of value and goal conflicts. Psychological Inquiry, vol.18, pp.1--22 (2007)
11. Kreps, D.M.: A Course in Microeconomic Theory. Princeton University Press, (1990).
12. Lotka A. J.: Elements of Physical Biology. Reprinted by Dover in 1956 as Elements of Mathematical Biology (1924)
13. Maynard S, J.: Models in Ecology. Cambridge University Press, Cambridge (1974)
14. Simon, H.A.: Behavioral Economics. In: The New Palgrave: A Dictionary of Economics. London, MacMillan (1987)
15. Tesfatsion, L.: Agent-based computational economics: Growing economies from the bottom up. Artificial Life, vol. 8, pp. 55--82 (2002)
16. Volterra, V.: Fluctuations in the abundance of a species considered mathematically. Nature vol. 118, pp. 558--560 (1926)

PART IV MODELING INVOLVEMENT IN ECONOMICAL CONTEXT

CHAPTER 8

A Personalized Intelligent Agent Model for Financial Decision Making Incorporating Psychological States and Characteristics for Greed and Risk

Part of this paper will appear as Bosse, T., Siddiqui, G.F., and Treur, J., Supporting Financial Decision Making by an Intelligent Agent Estimating Greed and Risk. In: Proceedings of the Fourth International Workshop on Human Aspects in Ambient Intelligence, HAI'10. IEEE Computer Society Press, 2010.

A Personalized Intelligent Agent Model for Financial Decision Making Incorporating Psychological States and Characteristics for Greed and Risk

Tibor Bosse[1], Ghazanfar F. Siddiqui[1,2], and Jan Treur[1]

[1] Vrije Universiteit Amsterdam, Department of Artificial Intelligence,
De Boelelaan 1081a, 1081 HV Amsterdam, The Netherlands
[2] Quaid-i-Azam University Islamabad, Department of Computer Science, 45320, Pakistan
{tbosse, ghazanfa, treur}@few.vu.nl ghazanfar@qau.edu.pk
http://www.few.vu.nl/~{tbosse, ghazanfa, treur}

Abstract. In the area of modeling financial decision making within economics it is more and more acknowledged that psychological states and personality characteristics play an important role. Examples of such states are feeling insecure in relation to financial risks, and being greedy in relation to opportunities to obtain serious gains. This paper presents an agent-based model of human decision making behavior in economic situations, based on such psychological states and characteristics. The model takes ideas underlying the Modern Portfolio Theory within economics into account, and incorporates psychological concepts like greed and a personality characteristic concerning risk. Thus a model is obtained that may provide a basis for the development of personalized intelligent agents that support a person in financial decisions. To evaluate the model a number of simulation experiments have been performed, which illustrate the model's ability to mimic investment behavior of different types of personalities. In addition, a mathematical analysis of equilibrium states of the model is presented.

1 Introduction

Financial decision making is not a trivial task for human beings. Already in 1979, Kahneman and Tversky [5] stated that people do not behave completely rationally when they have to decide between alternatives that involve risk, as, for example, in financial situations. Since then from time to time it has been argued that theories of economic decision making need to incorporate psychological factors such as greed and fear [3, 9, 11, 13].

The current paper is part of a project that aims to develop a personalized intelligent agent which supports persons that have to make investment decisions. The main goal of the project is to develop an agent that has insight in the individual psychological states and characteristics of persons that are working with financial applications, and is able to exploit this insight in order to provide appropriate support, both in following these states and characteristics in a personalized manner, and in encouraging reflection by the person through mirroring his or her states and decision making processes. For example, the agent may show the person how greedy he or she behaves.

In order to develop such a support agent, as a first step, a solid computational model of human decision making in financial context is needed. The development of such a model is the main contribution of the current paper. The model takes the main principles underlying the Modern Portfolio Theory (MPT) [4, 12] as a basis, and extends these with mechanisms to incorporate psychological factors (inspired, among others, by [2, 6, 8, 9, 11]).

The remainder of this paper is structured as follows. In Section 2, a global overview of the agent model for economic decision making is presented. In Section 3, this model is formalized in terms of mathematical formulae. Some simulations of the model are discussed in Section 4, and Section 5 presents a mathematical analysis of equilibrium states of the model. Finally, Section 6 provides a summary of the work and a discussion about future research.

2 Overview of the Agent Model

A global overview of the agent model for financial decision making and its interaction with the world is depicted in Figure 1. In this figure, the dotted boxes represent, from left to right, the human agent that makes the investment decisions, and the world. The small circles denote states of human and world. Note the difference between input states (depicted at the left hand side of a box), internal states (depicted within the box), and output states (depicted at the right hand side). The arrows indicate (causal) relationships between states.

As shown in Figure 1, the agent model includes a notion of greed as a mental state of an individual. This greed is assumed to be a dynamic state, which is continuously influenced by two other concepts, namely the observed results of investments and the individual's personality profile concerning risk taking or risk avoidance. The former is considered to be the result of (dynamic) real world investments, depending, among others on the state of the world's economy, whereas for the moment the latter is assumed to be a static characteristic of an individual.

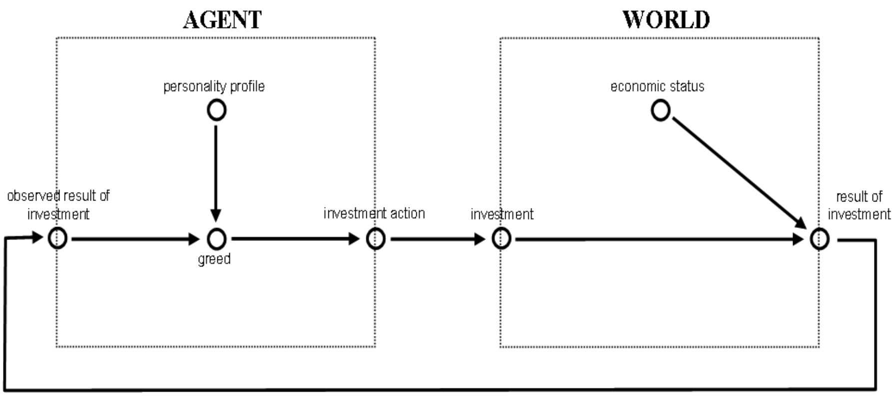


Figure 1. Overview of the Agent Model and its Interaction with the World.

Hence, the main idea of the model is that a person's greed is determined by her (long-term) personality profile (e.g., some persons are more risk seeking than others), combined with observations about recent events (e.g., if many investments have failed recently, persons are more likely to reduce their greed and as a consequence take less risk). This main idea is similar to the ideas in existing literature such as [2, 6, 8, 9, 11]. The greed directly affects the investment decisions that the individual makes: a person with higher greed will decide for more risky investments. To create an appropriate economic context, one particular type of investment decision is considered, namely the task to choose a financial product characterized by two factors concerning risk and expected gain. Thus the set of products is represented using a standard risk/return curve as proposed in the literature on Modern Portfolio Theory (see Section 3 for more details).

After an investment decision has been made, the selected product is transferred to the world, where it is determined what the result of the investment will be. As shown at the right hand side of Figure 1, this result depends not only on the selected product (in the sense that a more risky product has a lower probability to result in some return), but also on the economic state of the world. Thus, in case the situation of the economy becomes more positive, then the probability of receiving some return increases. Note that, for simplicity, the current model considers the economic status as an external variable, although in reality this variable may depend on many other factors, such as the economic behavior of other agents in the system (cf. [2]).

Finally, the results of investments are in turn observed by the individual, which completes the interactive loop between agent and world.

3 Formalizing the Agent Model

In this section, the global relations presented above are formalized in terms of mathematical specifications. In the model 10 products are used. Individuals are able to choose between these products by taking two characteristics into account, namely *risk* and *expected return*. To create a realistic range of products, the following parabolic equation is used for the relation between expected risk X and expected return Y (cf. [4, 12]):

$$X=aY^2+bY+c$$

with $a=1$, $b= -0.1$ and $c=0.1$.

The graph is shown in Figure 2. The idea is that more greedy persons will select products that are further to the right hand side of the curve. The formulae used in the model are as follows:

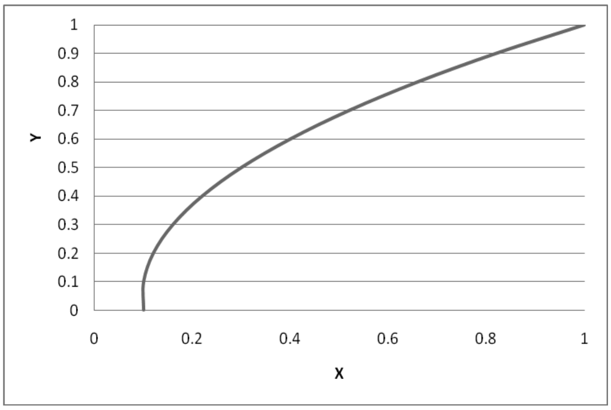


Figure 2. Expected Risk/Return Curve

Updating greed

The dynamics of greed are modeled by the following formula:

$$G(t+\Delta t) = G(t) + \beta\,(((p+1)/2)*E - G(t))\Delta t$$

In this formula, $G(t+\Delta t)$ is the updated greed, $G(t)$ is the old greed, E is the world event concerning the return on the earlier chosen product, and p is the individual static risk profile (0 means that the individual is low risk taking and 1 means the individual is high risk taking). The persistency factor β is the proportion of the old greed that is taken into account to determine the new greed. Simulation tests indicated that $\beta = 0.1$ give realistic results. The values of G, E and p are in the range between 0 and 1.

Thus, the underlying idea of the formula is that people may show more greedy behavior if their individual risk profile is more risk taking, and if they have received more positive experiences in the recent past (see also [2, 6, 8, 9, 11]).

Selecting a product

Next, based on the greed and the personality characteristic p the person selects a product. As a first step the following factor r is determined:

$$r = (1/G)-1\ ) / (2*(p+1))$$

This r is taken as the required slope of the curve (depicted in Figure 2) for the product to be chosen, according to Modern Portfolio Theory. The actual choice of the product is made as follows. For each of the considered products (X, Y) the following is calculated:

$$Z(X, Y) = Y - r*X$$

Then the product (X, Y) is chosen with the highest $Z(X, Y)$. This product is the closest approximation of the point at the curve with slope r.

Determining the return

The algorithm for calculating the return E of the investment is as follows; here W is the state of the world (taken between 0 and 1), and (X, Y) is the selected product:

1. *Generate a random number C between 0 and 1 (both inclusive)*
2. *IF $C \geq X * (1-W)$ THEN $E = Y$*
3. *IF $C < X * (1-W)$ THEN $E = 0$*

This shows, for example, that when the state of the economy *W* is maximal, there is no risk to have no return, and when *W* is minimal this risk is with probability *X*.

4 Simulation Results

Based on the agent model introduced in the previous section, a number of simulation experiments have been performed under different parameter settings of *p*. In Figures 3 to 11, the time is on the horizontal axis and *G, W, E* and the average profit received by the person are represented on the vertical axis. In all the simulation experiments, the initial value of greed is *0.5* and $\beta = 0.1$. For every simulation the value of *p* is different. For the value of *W*, a scenario has been established that is based on existing empirical data. For these data, the global Gross Domestic Product data over the period 1969-2008 have been taken from [16].

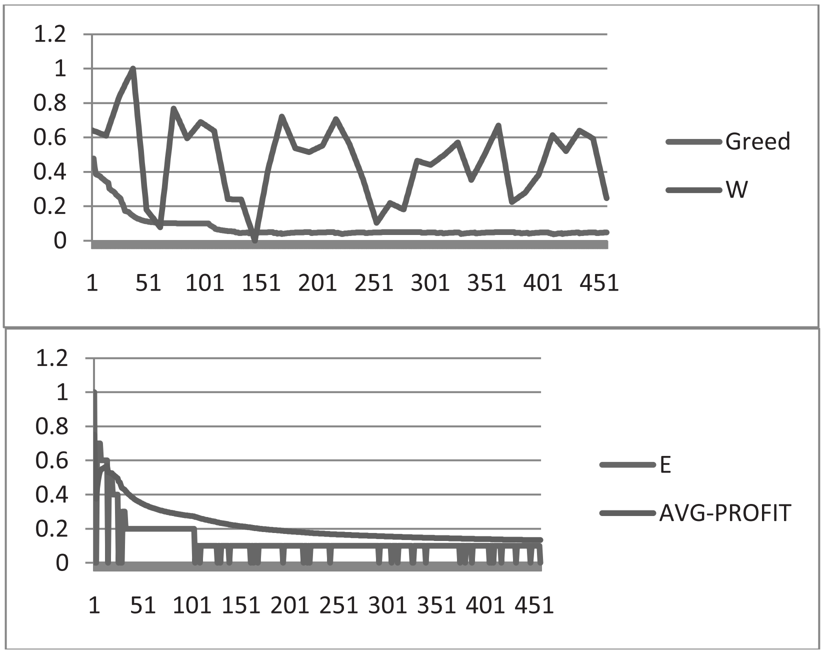


Figure 3. Simulation Results for p=0.

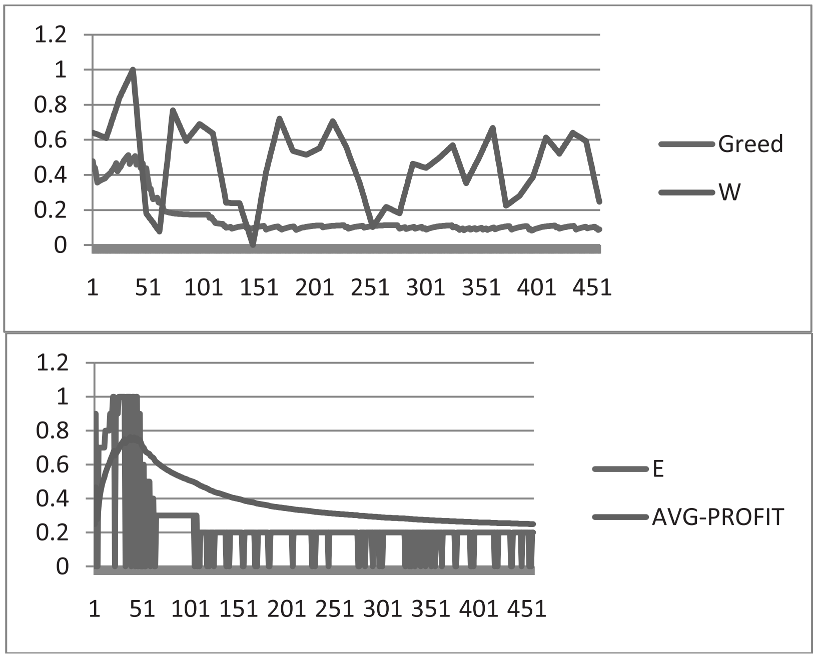


Figure 4. Simulation Results for p=0.15.

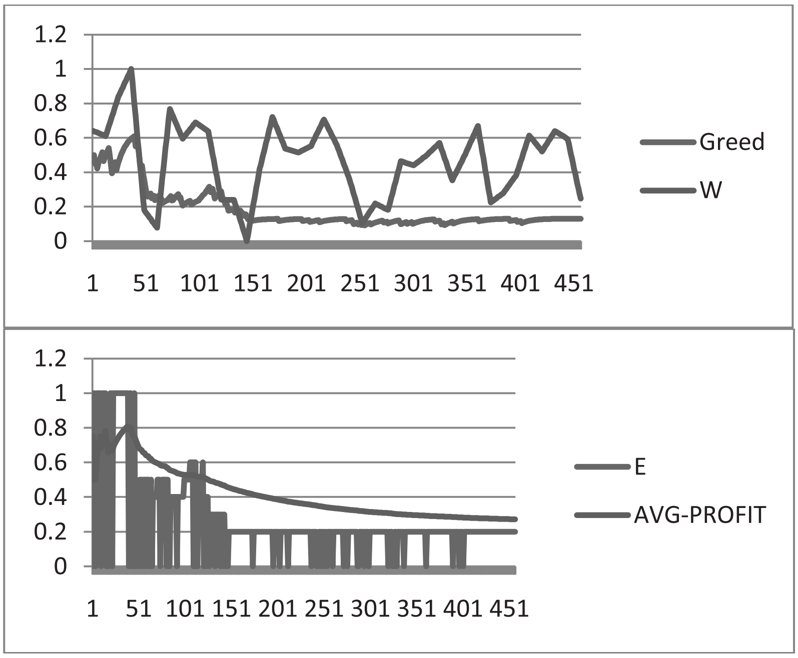


Figure 5. Simulation Results for p=0.3.

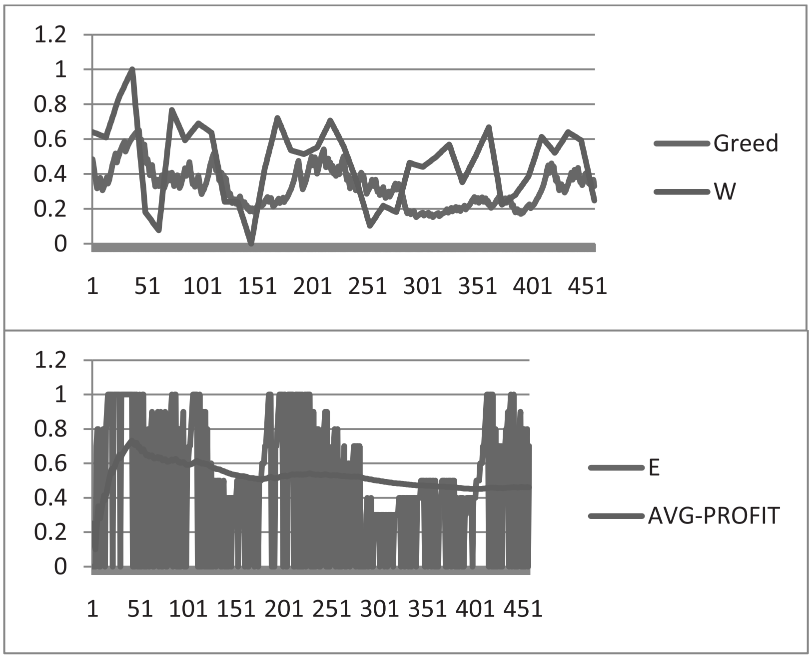


Figure 6. Simulation Results for p=0.4.

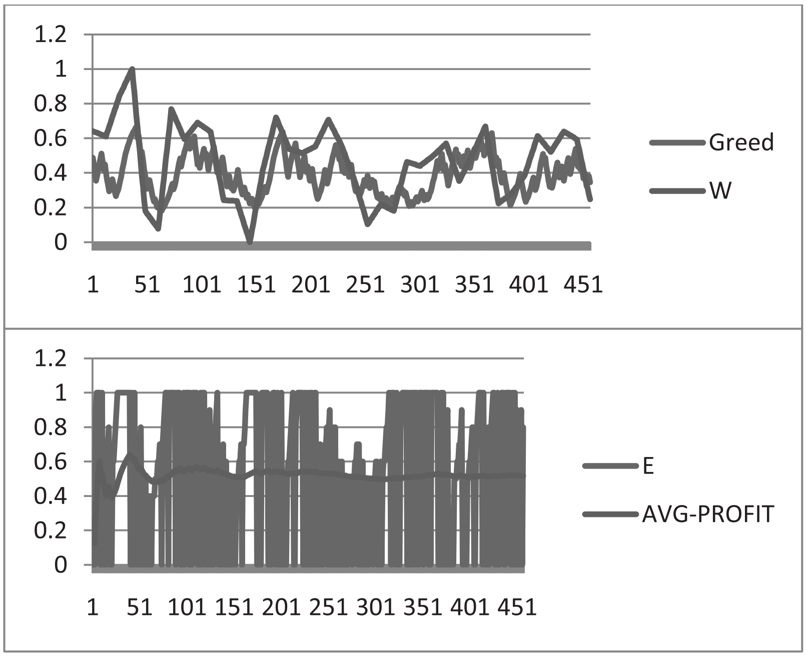


Figure 7. Simulation Results for p=0.5.

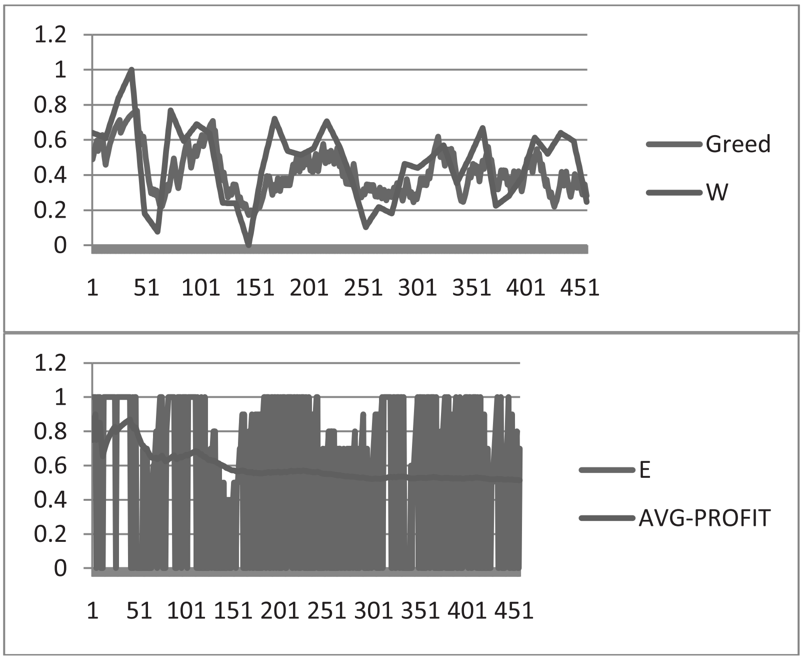


Figure 8. Simulation Results for p=0.6.

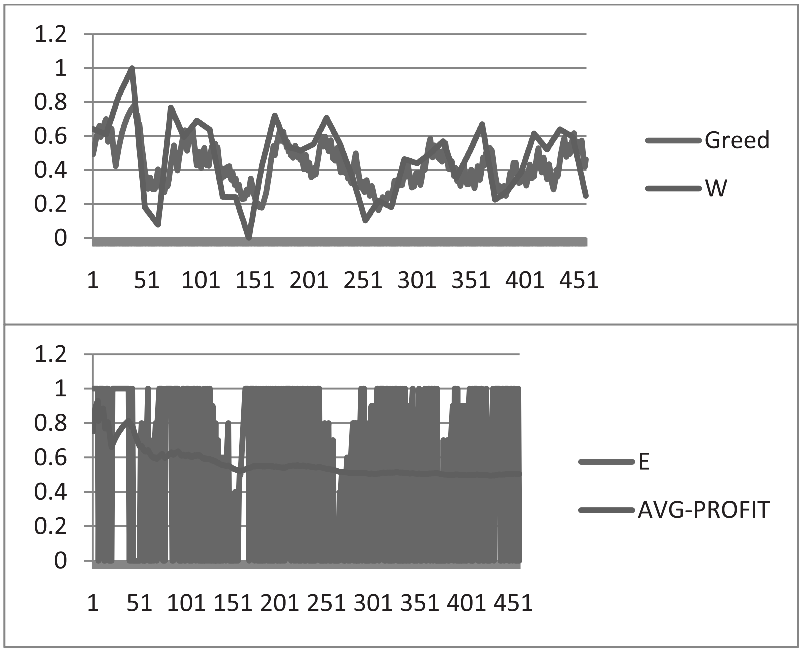


Figure 9. Simulation Results for p=0.7.

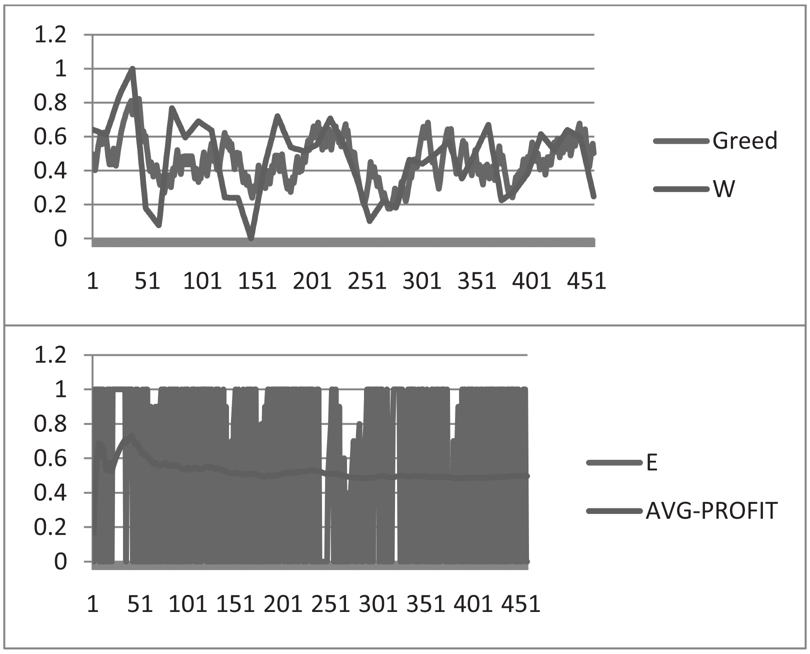


Figure 10. Simulation Results for p=0.85.

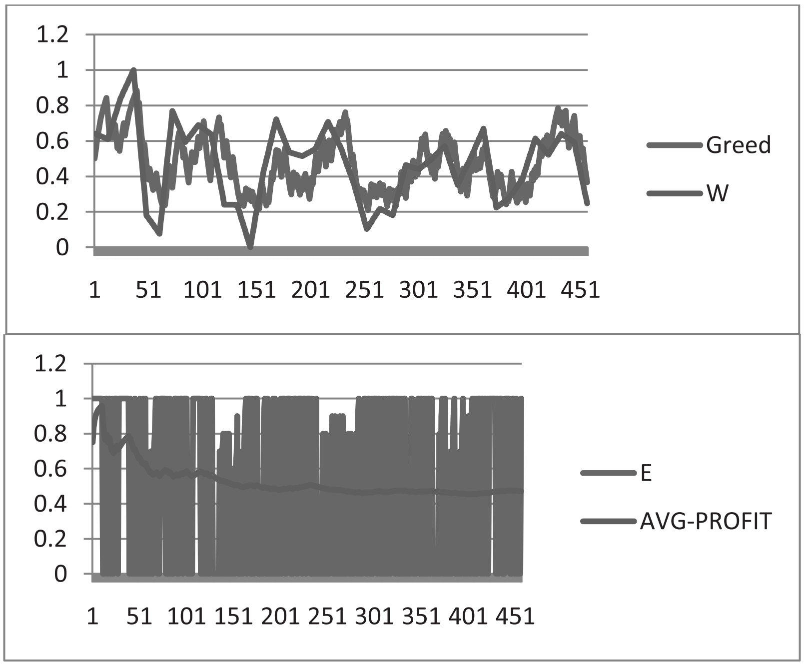


Figure 11. Simulation Results for p=1.

As illustrated by these simulation, in most cases the person adapts its greed *G* to the status of the economy *W*. Thus, the persons quickly learn which level of greed is most appropriate in which situation. Only in the cases with very low *p*, after a while the greed becomes so low that the person cannot recover from that anymore (because it will never have very positive experiences).

Moreover, when comparing the different graphs with each other, one can note that the values of the events (and thus also the average profit) are highest in the case where *p=0.6*. In cases with a higher *p*, the average profit slightly decreases, although not much.

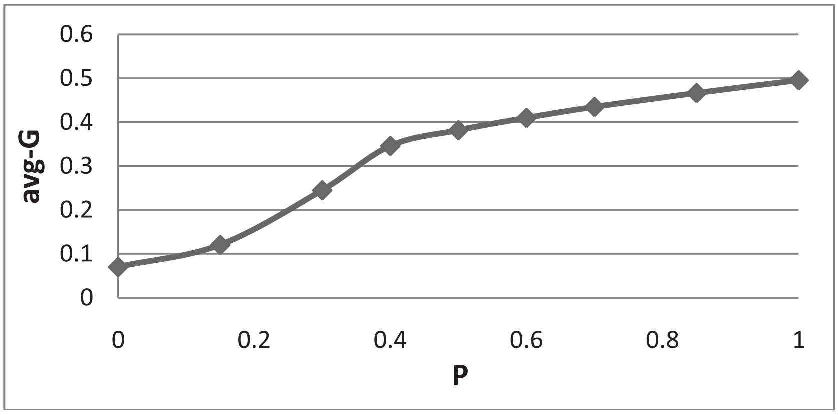


Figure 12. Relation between average greed and p.

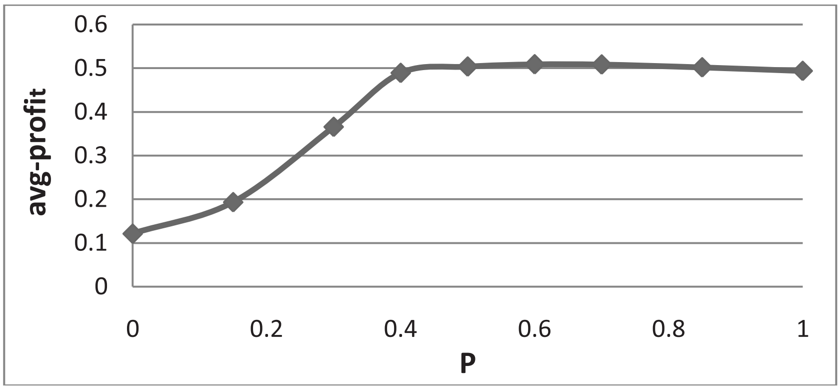


Figure 13. Relation between profit and p.

In addition, a larger number of simulation runs has been performed to see the effect of *p* on the average greed and average profit over the simulation. For this, 50 simulations were taken for each value of *p*, and graphs were plotted between *p*, average greed and average profit. In Figure 12, and Figure 13, *p* is on the horizontal axis and average greed or average profit is on the vertical axis. It is evident from Figure 12 that as *p* increases, the average greed also increases.

Furthermore, Figure 13 shows roughly that the average profit increases with an increase in the value of *p*. The average profit increases sharply from *p=0* to *p=0.4*, and then becomes more or less stable. The point where average profit is maximal lies at $p=0.6$, as also predicted based on the simulation shown in Figure 8. Hence, apparently this is a scenario where it is optimal to have a personal risk profile of 0.6. In other scenarios (e.g., with a very bad economy), other risk profiles may be more beneficial.

5 Mathematical Analysis

In this section a mathematical analysis of equilibrium states of the model is discussed. For any given fixed values of p and W, the process indeed approaches an equilibrium state, as was found by various example simulations. Knowledge of how these equilibrium states are for each pair of values for of p and W, is also useful to analyze functioning of the process for changing W, as after each change in W, the process will aim to reach an equilibrium state for this new value.

The general setup is based on a continuous model given by the following differential equation for G:

$$dG/dt = \beta\, (((p+1)/2)E - G)$$

The products characterized by pairs (X, Y) are taken from the curve

$$X = a\,Y^2 + bY + c$$
$$\text{with} \quad a=1 \quad b=-0.1 \quad c = 0.1$$

Each time the product is chosen on this curve for which $dY/dX = r$ or $dX/dY = 1/r$ (according to Modern Portfolio Theory) with $r=((1/G)-1)/(2(p+1))$.

Relation between *G* and the product chosen

From the curve equation follows that $dX/dY = 2aY + b$. Therefore:

$$2aY + b = 1/r = 2(p+1)/((1/G)-1) =$$
$$2(p+1)G/(1-G)= 2(p+1)G/(1-G)$$

This can be used to express y in G as follows:

$$2aY = 2(p+1)G/(1-G) - b$$

$$\begin{aligned} Y &= (2(p+1)G/(1-G) - b)/2a \\ &= (2(p+1)G - b(1-G))/2a(1-G) \\ &= ((2p+2+b)G - b)/2a(1-G) \\ &= ((2p+1.9)G +0.1)/2(1-G) \\ &= ((20p+19)G + 1)/20(1-G) \end{aligned}$$

So, at each point in time

$$Y = ((20p+19)G + 1)/20(1-G)$$
$$X = Y^2 -0.1Y + 0.1$$

Some special cases are:

$$G = 0 \;\Rightarrow\; Y = 1/20$$
$$p = 0 \;\Rightarrow\; Y = (19G + 1)/20(1-G)$$

Determining an expectation value for E

An equilibrium for G has to satisfy $dG/dt = 0$. Given the differential equation this is equivalent to $(((p+1)/2)E = G$. However, the events that are considered outcomes E depend on a random number C in $[0, 1]$: if $C \geq X.(1-W)$ then $E=Y$, else $E=0$. Therefore:

$E = 0$ with probability $X.(1-W)$
$E = Y$ with probability $1 - X.(1-W)$

From this the expectation value $ExpVal(E)$ for E is determined as follows:

$$
\begin{aligned}
ExpVal(E) &= X.(1-W) *0 + (1 - X.(1-W)) *Y \\
&= (1 - X.(1-W))\, Y
\end{aligned}
$$

As E itself is always fluctuating due to the randomness no real equilibrium can occur. Therefore, for an equilibrium analysis it is better to take the expectation value $ExpVal(E)$ instead of E itself.

The equilibrium equations for G

An estimation for an equilibrium can be obtained when for E its expectation value is taken. Then the equation for $dG/dt = 0$ is:

$$(((p+1)/2)*ExpVal(E) = G$$

This can be rewritten into

$$(((p+1)/2)* (1 - X.(1-W))\, Y = G$$

$$G = 0.5\, Y\, (p+1)(1 - X.(1-W))$$

This equation can be combined with the previous one to form the following set of three equilibrium equations with three unknowns:

$$
\begin{aligned}
&\mathbf{G = 0.5\, Y\, (p+1)(1 - X.(1-W))} \\
&\mathbf{X = Y^2 - 0.1Y + 0.1} \\
&\mathbf{Y = ((20p+19)G + 1)/20(1-G)}
\end{aligned}
$$

An example solution obtained from simulations for $p=1$ and $W=0$ is:

$G = 0.226$ $Y = 0.265$ $X = 0.144$

Some special cases are:

$p = 0 \Rightarrow G = 0.5\, Y\, (1 - X.(1-W))$
$p = 1 \Rightarrow G = Y\, (1 - X.(1-W))$
$W = 1 \Rightarrow G = 0.5(p+1)Y$
$p = 0\ \&\ W = 1 \Rightarrow G = 0.5\, Y$
$p = 1\ \&\ W = 1 \Rightarrow G = Y$

Overview of solutions of the equilibrium equations

A large number of simulations have been performed for different values of *p* and *W*. The outcome of the equilibrium values for G found is shown in Figures 14 (*G* against *p* for different values of *W*) and 15 (*G* against *W* for different values of *p*).

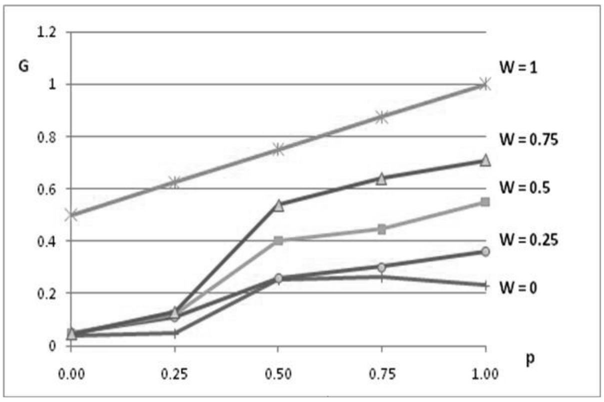


Figure 14. Dependence of the equilibrium value of G on p for some values of W.

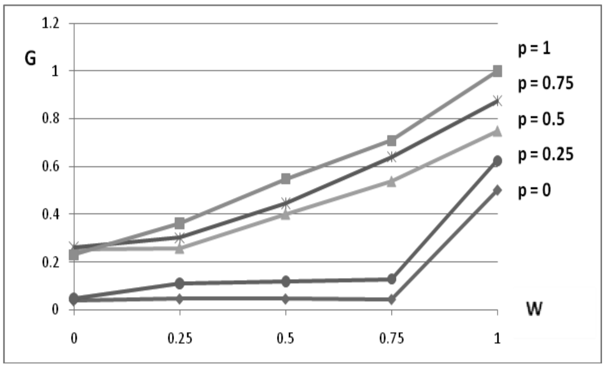


Figure 15. Dependence of the equilibrium value of G on W for some values of p.

In Figure 14 it is shown that practically always the equilibria values for *G* depend on the basic personal characteristic *p* in a monotonically increasing manner; this expresses that the personal characteristic *p* indeed relates to the greed of the person, even while the latter is dynamic. Moreover, in Figure 15 it is shown that the equilibria values for *G* also depend on the state *W* of the economy in a monotonically increasing manner. This shows that a better economy leads to higher greed. Note, however, that for low values of *p* (from *0* to *0.3*) the state *W* of the economy has not much impact, unless it is extremely high (*W* close to *1*).

Solving the equilibrium equations

In principle, the equilibrium equations can be solved symbolically by algebraic manipulation in order to obtain an algebraic expression for *G* as a function of *p* and *W*. The following equation for *G* can be obtained:

$$(p+1)(1 - X.(1-W))\ Y = 2G$$

$$(p+1)(1 - (Y^2 - 0.1Y + 0.1).(1-W))\ Y = 2G$$

$$(p+1)(1 - ((((20p+19)G\ + 1)/20(1-G))^2 -$$
$$0.1(((20p+19)G + 1)/20(1-G)) + 0.1).$$
$$(1-W))\ (((20p+19)G\ + 1)/20(1-G)) = 2G$$

$$(p+1)(1 - ((((20p+19)G\ + 1)^2 -$$
$$0.1(((20p+19)G\ + 1)\ 20(1-G)) + 0.1(20(1-G))^2).$$
$$(1-W))\ (((20p+19)G\ + 1) = 2G\ (20(1-G))^3$$

Although this fourth degree equation in *G* can be solved symbolically to obtain *G* as a function of *p* and *W*, this is omitted as it provides a rather complex expression for *G* as a function of *p* and *W*. A second way of solving the equilibrium equations is by numerical approximation methods such as Newton-Raphson [7]. This can be only done for given values of *p* and *W*.

A third way of solving the equilibrium equations is to derive a differential equation for *G* as a function of *p*. This can be done by differentiating the three equilibrium equations to *p* as follows:

$$dG/dp = d0.5\ Y\ (p+1)(1 - X.(1-W))\ /dp$$

$$d20(1-G)Y/dp = d\ ((20p+19)G\ + 1)\ /dp$$

$$dX/dp = d(Y^2 - 0.1Y + 0.1)\ /dp$$

thus obtaining:

$$(p+1)\ (1 - X.(1-W))\ \mathbf{dY/dp} + 2G/(p+1) =$$
$$2\ \mathbf{dG/dp} + (p+1)\ Y.(1-W)\ \mathbf{dX\ /dp}$$

$$20(1-G)\ \mathbf{dY/dp} = (20p + 20Y + 19)\ \mathbf{dG/dp} + 20G$$

$$\mathbf{dX/dp} = (2Y - 0.1)\ \mathbf{dY/dp}$$

Here the third equation can be used to eliminate ***dX/dp*** from the first two equations, providing two equations in ***dG/dp*** and ***dY/dp***, which can be solved by expressing each of ***dG/dp*** and ***dY/dp*** in *G, X, Y* (and *p* and *W*). From these *X* can be easily eliminated by using the curve equation. The resulting two first-order differential equations for *G* and *Y* can be solved numerically for any given starting point. It is also possible to derive one single first-order differential equation for *G* by replacing *Y* by *((20p+19)G+1) / 20(1-G)* according to the third equilibrium equation. This single differential equation can also be used for numerical simulation.

6 Discussion

This paper presents an agent-based model of human decision making behavior in economic situations, based on psychological states and characteristics concerning greed and risk taking or risk avoidance. The model takes ideas underlying the Modern Portfolio Theory [4, 12] within economics into account, and incorporates the psychological concept greed and a risk characteristic. Thus a model is obtained that may provide a basis for the development of personalized intelligent agents that support a person in financial decisions.

In the area of modeling financial decision making within economics it is more and more acknowledged that psychological states and personality characteristics play an important role [3, 9, 11, 13]. Examples of such states are feeling insecure in relation to financial risks, and being greedy in relation to opportunities to obtain serious gains. The proposed agent model provides a computational formalization of such concepts.

To evaluate the model a number of simulation experiments have been performed, which illustrate the model's ability to mimic investment behavior depending on the types of personality and the state of the economy. A mathematical analysis was contributed of the equilibrium states of the model.

In recent years, various authors have studied processes in economics using agent-based simulation techniques (e.g., [2, 14, 15]. However, the current paper differences from these approaches in that it attempts to study the decision making behavior of a single agent in detail, instead of analyzing the global dynamics of a multi-agent system. Future work aims at validation of the model, and at development of a virtual agent (avatar) [10] that learns the person's personality profile and supports the person by indicating choices of products that fit to the person's states and personality, but also by mirroring these aspects and implied behavior in order to stimulate reflection by the person. As pointed out by various authors (e.g., [1]), virtual characters with personalization features can augment involvement of users in financial (e.g., banking) applications.

References

1. Blom, J. and Monk, A. (2001). One-to-one e-commerce: who's the one? In: Proceedings of the Conference on Human Factors in Computing Systems (CHI'01). ACM Press, pp. 341-342.
2. Bosse, T., Siddiqui, G.F., and Treur, J. (2010). Modeling Greed of Agents in Economical Context. In: Proceedings of the 23rd International Conference on Industrial, Engineering & Other Applications of Applied Intelligent Systems, IEA/AIE'10. Lecture Notes in Artificial Intelligence, Springer Verlag, 2010, in press.
3. Brandstätter, H. (1993). Should Economic Psychology Care about Personality Structure? Journal of Economic Psychology, vol.14, pp. 473-494.
4. Haugen, R.A. (1997). Modern Investment Theory. Prentice-Hall, Inc.
5. Kahneman, D. and Tversky, A. (1979). Prospect Theory: An Analysis of Decision under Risk. Econometrica, XLVII, pp. 263-291.
6. Kasser, T., Cohn, S., Kanner, A., Ryan, R. (2007). Some costs of American corporate capitalism: A psychological exploration of value and goal conflicts. Psychological Inquiry, vol.18, pp.1-22.
7. Kelley, C.T. (2003). Solving Nonlinear Equations with Newton's Method. Fundamentals of Algorithms, vol. 1. SIAM, Philadelphia, PA.
8. Krawiec, K.D. (2000). Accounting for Greed: Unravelling the Rogue Trader Mystery. Oregon Law Review, vol 79, issue 2, pp. 301-339.
9. Lo, A., Repin, D.V., and Steenbarger, B.N. (2005). Fear and Greed in Financial Markets: A Clinical Study of Day-Traders. American Economic Review 95, pp. 352-359.
10. Prendinger, H., Lester, J., and Ishizuka, M. (eds.) (2008). Intelligent Virtual Agents. Proceedings of the 8th International Conference on Intelligent Virtual Agents, IVA'08. Springer LNAI, vol. 5208.

11. Rabin, M. (2002). A Perspective on Psychology and Economics. European Economic Review, vol. 46, pp. 657-685.
12. Sabal, J. (2002). Financial Decisions in Emerging Markets. Oxford University Press, Inc., New York.
13. Simon, H.A. (1987). Behavioral Economics. In: The New Palgrave: A Dictionary of Economics. London, MacMillan.
14. Terano, T., Deguchi, H., and Takadama, K. (eds.) (2003). Meeting the Challenge of Social Problems via Agent-Based Simulation. Post-Proceedings of the 2nd International Workshop on Agent-Based Approaches in Economic and Social Complex Systems. Springer Verlag, 2003.
15. Tesfatsion, L. (2002). Agent-based computational economics: Growing economies from the bottom up. Artificial Life, vol. 8, pp. 55-82.
16. http://www.ers.usda.gov/Data/Macroeconomics/

PART V EMBODYING EMOTIONS IN VIRTUAL AGENTS

CHAPTER 9

Incorporating Emotion Regulation into Virtual Stories

This chapter appeared as Bosse, T., Pontier, M., Siddiqui, G.F., and Treur, J., Incorporating Emotion Regulation into Virtual Stories. In: Pelachaud, C., Martin, J.C., Andre, E., Chollet, G., Karpouzis, K., and Pele, D. (eds.), Proceedings of the Seventh International Conference on Intelligent Virtual Agents, IVA'07. Lecture Notes in Artificial Intelligence, vol. 4722. Springer Verlag, 2007, pp. 339-347.

Incorporating Emotion Regulation into Virtual Stories

Tibor Bosse, Matthijs Pontier, Ghazanfar F. Siddiqui, and Jan Treur

Vrije Universiteit Amsterdam, Department of Artificial Intelligence,
De Boelelaan 1081a, 1081 HV Amsterdam, The Netherlands
{tbosse, mpontier, ghazanfa, treur}@few.vu.nl
http://www.few.vu.nl/~{tbosse, mpontier, ghazanfa, treur}

Abstract. This paper presents an approach to incorporate emotion regulation as addressed within psychology literature into virtual characters. To this end, first Gross' informal theory of emotion regulation has been formalized using a dynamical system style modeling approach. Next, a virtual environment has been created, involving a number of virtual agents, which have been equipped with the formalized model for emotion regulation. This environment has been used to successfully generate a number of emergent virtual stories, in which characters regulate their emotions by applying regulation strategies such as situation selection and attentional deployment. The behaviors shown in the stories were found consistent with descriptions of human regulation processes.

1 Introduction

In recent years, there has been an increasing interest in the area of *virtual storytelling*, addressing the development of computer systems that generate fictive stories in which the characters show realistic behavior. In order to develop virtual stories, a large variety of approaches have been proposed, e.g., [4], [5], [13]. A trend that can be observed in many of these approaches is the movement from stories with a fixed, pre-scripted storyline towards emergent narrative, i.e., stories in which only a number of characters and their personalities are fixed, rather than the precise script of the story [1]. In the latter type of storytelling, ideally, all the designer (or writer) has to do is to determine which (types of) characters will occur in the play (although usually it is still needed to roughly prescribe the course of events). Hence, advantages of emergent narrative are the reduced amount of work that has to be spent by the writer, and the non-deterministic and unpredictable behavior of the story.

In parallel with the shift from fixed storylines to emergent narrative, there has been a development in the nature of the involved characters as well. Recently, the characters (or agents) that are present in virtual stories are transforming more and more from shallow avatars to complex personalities with human-like properties such as emotions and theories of mind, e.g., [15]. To accomplish this, researchers have started to incorporate cognitive models within virtual characters, e.g., [10], [12]. Despite these first promising attempts, there is still a wide area to explore when it comes to enhancing virtual agents with cognitive capabilities.

In line with the development described above, this paper explores the possibilities to equip the characters involved in virtual stories with the capability of *emotion regulation*. Informally, emotion regulation can be described as the process humans undertake to increase, maintain or decrease their emotional response, see e.g., [7], [8], [11], [14]. The idea is that, by offering virtual agents the capacity to actively regulate their emotions, they will be able to select those kinds of behaviors that they feel most comfortable with. As a result, such agents will 1) behave more realistically and 2) have more freedom in the choice of their actions, which enhances the emergent narrative effect. This approach is similar to the approach taken in [9], which aims at incorporating coping behavior into virtual humans.

In order to build emotion regulation into virtual stories, in this paper the informal model by Gross [7] as found in psychology literature was taken as a basis. This model describes a number of strategies humans use to adapt their emotion response levels, varying from situation selection to cognitive change and response modulation. Next, this model has been formalized using a dynamical system style modeling approach (see also [3] for some initial steps). In addition, a virtual environment has been created, incorporating a number of virtual agents, and these agents have been equipped with the formalized model for emotion regulation. To test the behavior of the model, a series of simulation experiments has been performed using the LEADSTO simulation language [2]. The model has been connected to the Vizard Virtual Reality Toolkit [16], to visualize the resulting stories in a graphical environment.

2 Emotion Regulation in the Virtual Agent Context

Gross [8] describes a process model of emotion regulation using the following definition: 'Emotion regulation includes all of the conscious and nonconscious strategies we use to increase, maintain, or decrease one or more components of an emotional response'. In his model, Gross distinguishes four different types of emotion regulation strategies, which can be applied at different points in the process of emotion generation. First of all, when applying *situation selection*, a person chooses to be in a situation that matches the emotional response level the person wants to have for a certain emotion. For example, you can stay home instead of going to a party, because you are in conflict with someone who is going to that party. Second, when applying *situation modification*, a person modifies an existing situation so as to obtain a different level of emotion. For instance, when watching an irritating television program, you zap to another channel. Third, *attentional deployment* refers to shifting your attention to a certain aspect. For example, you close your eyes when watching an exciting penalty shoot-out. Finally, *cognitive change* refers to selecting a cognitive meaning to an event. For example, when a person loses a tennis match and blames the weather circumstances, instead of his own capacities.

To incorporate these strategies into virtual characters, a modeling approach was used that is based on the LEADSTO simulation environment [2] and the Vizard Virtual Reality Toolkit [16]. Due to space limitations, the technical details of LEADSTO and Vizard are not shown here. However, they can be found in Appendix A in [17]. Below, in Section 2.1, at a language-independent level a global overview is given of the model, of which an initial version can be found in [3]. Next, in Section 2.2, for each of the regulation strategies it is shown how it is used in the virtual agents playing as characters in virtual stories. The complete formal specification of the model (in LEADSTO notation) is shown in Appendix B in [17].

2.1 Global Overview

In order to incorporate emotion regulation strategies into virtual agents, a virtual environment is created that is populated by a number of agents. Each agent is equipped with a mechanism to regulate its emotions, which is based on the model as described informally by Gross [7]. To create a formal model, for any given type of emotion a number of variables have been introduced. For convenience, the model concentrates on one specific type of emotion. In principle, this can (at least) be any emotion that is considered to be a basic human emotion, e.g., sadness, happiness, or anger [6]. In order to describe the regulation of such an emotion, the model takes into account the four strategies discussed by Gross are used (i.e., *situation selection*, *situation modification, attentional deployment*, and *cognitive change*). Based on the four strategies mentioned, in the formalization four corresponding *elements* (denoted by k) are introduced, for the objects that are affected by the particular strategies: *situation*, *sub-situation*, *aspect*, and *meaning*.

The model assumes that each agent aims at an optimal level of emotion. The regulation process in the virtual agents starts by comparing the actual *emotion response level* ERL to the emotion response level ERL_norm aimed at. The difference between the two is the basis for adjustment of the choices made for each of the elements k; based on these adjusted choices, each element k will provide an adjusted *emotional value* EV_k.

To obtain a quantitative model, the emotion response level and the emotional values for the different elements for a given type of emotion are represented by real numbers in the interval [0, 2] (where 0 is the lowest possible ERL (e.g., extreme sadness), and 2 the highest (e.g., extreme happiness)). In the model, the level of emotion to aim at (the ERL norm), is also expressed in a real number in the domain [0, 2]. Based on these concepts, the ERL is recalculated each step by the following difference equation formula:

$$new\_ERL = (1-\beta) * \Sigma_k (w_k * EV_k) + \beta * ERL$$

In this formula, new_ERL is the new emotion response level, and ERL is the old emotion response level. The persistency factor β is the proportion of the old emotion response level that is taken into account to determine the new emotion response level. Initial tests have indicated that values for β around 0.9 deliver realistic results. The new contribution to the emotion response level is calculated by the weighted sum of the emotional values: $\Sigma_k\ w_k * EV_k$. By normalization, the sum of all the weights w_k is taken to be 1. The following section describes how the different strategies influence the values of EV_k.

2.2 Emotion Regulation Strategies

Situation selection: which agent to meet. Every step, each agent chooses to be alone, or to contact another agent, by comparing the EVs it attaches to being alone and to being with other agents. The agent will always choose the option with the EV that is closest to its optimal ERL. When two agents contact each other, they decide to meet. When the agents are meeting, their EV for situation is set to the EV they attach to the other agent. When an agent chooses to be alone, its EV for situation is set to its EV for being alone.

Situation modification: what to talk about. When two agents are in a meeting, they will talk about a certain conversation subject. To decide which of the agents will start talking, each agent has a personal *dominance factor*. The agent with the highest dominance factor will choose the first conversation subject. Each step after this, the agent who has not chosen the current conversation subject will choose the next conversation subject. When an agent gets to choose which conversation subject to talk about, it will compare the EVs it attaches to each conversation subject, and select the one that is closest to its optimal ERL. The EV for subsituation is set to the EV the agents attach to the conversation subject they are currently talking about. When an agent is not in a meeting, its EV for subsituation will be set to the neutral value of 1, since the agent is not in a subsituation. When an agent A talks to another agent B about a certain conversation subject CS, this will affect the way agent B thinks about agent A. Agent B's EV for agent A will change using the following formula:

$$new\_EV_{agent\_A} = \beta_{friendship} * EV_{agent\_A} + (1-\beta_{friendship}) * EV_{CS}$$

In this formula, $new\_EV_{agent\_A}$ is the new EV agent B will attach to agent A and EV_{agent_A} is the old EV agent B attached to agent A. The persistency factor $\beta_{friendship}$ is the proportion of the old EV that is taken into account to determine the new EV. Here, values for $\beta_{friendship}$ bigger than 0.9 (where $\beta_{friendship}$ will get bigger when an agent knows another agent for a longer time) deliver realistic results. The new contribution to the ERL is determined by EV_{CS}: the EV agent B attaches to the conversation subject agent A is talking about. So how much an agent likes another agent, depends on how much an agent likes the conversation subjects another agent talks about.

The extent to which an agent likes to talk about a certain conversation subject can be changed by external events. For example, an agent will start to like a sports team more when this team wins a match. To accomplish this, the following formulas are used:

$$new\_EV_{CSn} = EV_{CSn} + \Delta EV_{CSn}$$

When a positive event occurs: $\Delta EV_{CSn} = \eta * EV_{CSn} * (d_{max} – EV_{CSn})$

When a negative event occurs: $\Delta EV_{CSn} = -\eta * EV_{CSn} * (d_{max} – EV_{CSn})$

In these formulas, new_EV_{CSn} is the new EV the agent attaches to the conversation subject, and EV_{CSn} is the old EV the agent attached to the conversation subject. Here η is a variable that determines the speed of adjusting EVs to conversation subjects. A lower η will result in slower adjustment. Here, an η of 0.02 delivers realistic results. The part $EV_{CSn} * (d_{max} – EV_{CSn})$ prevents EV_{CSn} from under- or overadjustment.

Attentional deployment: on which aspect to focus. When an agent is in a conversation, it can choose to pay attention to, or to distract its attention from the conversation. Every step, the agent chooses the option with the EV closest to its optimal ERL. The EVs the agent attaches to paying attention or distracting its attention, depends on the conversation subject the agent is currently talking about, according to the following formulas:

$$new\_EV_{pay\_attention} = \beta_{asp} * EV_{pay\_attention} + (1-\beta_{asp}) * EV_{CS}$$

$$new\_EV_{distract} = \beta_{asp} * EV_{distract} + (1-\beta_{asp}) * (-EV_{CS} + d_{max})$$

In these formulas, $new\_EV_{pay\_attention}$ and $new\_EV_{distract}$ are the new EVs for pay_attention and distract, and $EV_{pay\_attention}$ and $EV_{distract}$ are the old EVs for pay_attention and distract. The persistency factor β_{asp} is the proportion of the old EV that is taken into account to determine the new EV. The new contribution to the EV for pay_attention is determined by EV_{CS}, the EV the agent attaches to the conversation subject it is talking about. The new contribution to the EV for distract is calculated by $(-EV_{CS} + d_{max})$. This will reach a high value when the agent attaches a low EV to the conversation subject, and a low value when the agent attaches a high value to the conversation subject. So when the agent likes the conversation subject, it will be more likely to pay attention to the conversation. The agent chooses to distract from, or pay attention to the conversation, by comparing the two EVs for paying attention and distracting, and picking the option with the EV closest to its optimal ERL.

Cognitive change: which meaning to attach. Every step, agents can choose to apply self-talk. An agent can use self-talk to relativize its current state of mind, or on the other hand, to attach more meaning to its current state. Every step, an agent chooses to relativize, attach a stronger meaning, or to apply no self-talk, by picking the option with the EV closest to the optimal ERL of the agent. The EV for not applying self-talk always has the neutral value of 1. The EVs for relativizing and attaching more meaning depend on the ERL of the agent, and are updated every step according to the following formula's:

$$new\_EV_{relativise} = d_{max} - ERL$$
$$new\_EV_{attach\_more\_meaning} = ERL + (ERL-1) * (1 - abs(1-ERL))$$

When an agent has a high ERL, the EV for relativising will be low, and when an agent has a high ERL, the EV for relativizing will be high. So relativizing always influences the ERL of the agent to reach a more neutral value.

When the ERL of the agent has the neutral value of 1, (ERL-1) will be 0, and the EV for attaching more meaning will be 1. When the ERL of the agent is smaller than 1, then ERL-1 will have a negative value, and the EV for attaching more meaning will have a value that is smaller than the current ERL. When the ERL of the agent is bigger than 1, then ERL-1 will be bigger than 1, and the EV for attaching more meaning will have a value that is bigger than the current ERL. So attaching more meaning always influences the ERL of the agent to a more extreme value than the current one. Multiplying by (1 – abs(1-ERL)) prevents the EV from reaching values that are out of the domain.

3 Simulation Experiments

Several experiments have been done to test the simulation model's ability to generate interesting scenarios. To obtain movies in Vizard, events in the LEADSTO simulations were translated to visualizations in Vizard. The exact mapping that was used for this translation is shown in Appendix C in [17]. For example, the fact that an agent is happy is visualized by a certain type of smile, and the fact that an agent distracts from a conversation is visualized by this agent moving its head away from its conversation partner.

In all of the simulations, three agents are involved, which will be called Barry, Gary, and Harry. The particular emotion these agents will try to regulate during the scenario's is their amount of happiness. To enable this, the particular topics they are allowed to talk about are football (in particular, the Dutch football teams Ajax and Feyenoord) and hockey. The parameter settings of all agents used in three specific experiments are shown in Appendix D in [17].

Due to space limitations, only one of the simulation experiments is discussed in this paper. The results of the LEADSTO simulation of this experiment can be seen in Figure 1. Here, time is on the horizontal axis, whereas different events are displayed on the vertical axis. A dark box on top of a line indicates that an event is true at that time point; a light box below a line indicates than an event if false. A detailed description of what happens in this scenario is provided in Appendix E in [17].

As mentioned earlier, using a specific conversion program that has been implemented, LEADSTO simulations were translated into movies in Vizard. A screenshot of an example Vizard movie (which corresponds to the scenario shown in Figure 1) is shown in Figure 2. This figure shows a situation in which (on the foreground) two agents are having a conversation. The left agent is talking about hockey, but the right agent tries to distract from the conversation by moving its head away from the conversation. The cognitive meaning that each agent attaches to its current thoughts is displayed (in red) above the heads of the agents. Meanwhile, in the background a third agent is standing alone. The full Vizard movie of this scenario (as well as the movies that correspond to the two other experiments described in Appendix D) can be found on [17].

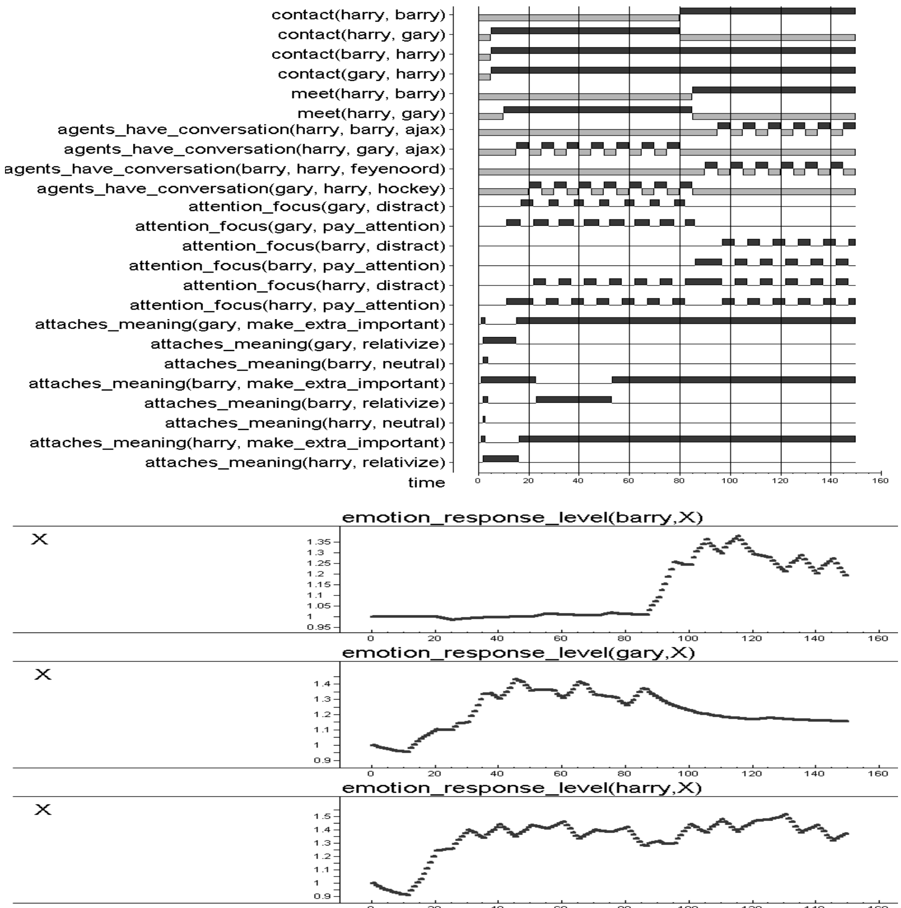


Fig. 1. Example Simulation Trace

The resulting movies provide a first indication that the emotion regulation strategies as described by [7] have been implemented successfully within the virtual agents used as characters. To be specific, the agents are able to perform *situation selection* by selecting different conversation partners, and withdrawing from conversations if desired. Moreover, they can perform *situation modification* by changing conversation topics, they can perform *attentional deployment* by changing the amount of attention they pay to a conversation, and they can perform *cognitive change* by changing the cognitive meaning they assign to their thoughts (e.g., by stating to themselves that something is not very important). These behaviors were found consistent with predicted behaviors for situations as described by Gross [7], [8] (which are based on empirical evidence).

Fig. 2. Screenshot of an example scenario in Vizard

4 Discussion

Within the domain of virtual storytelling, the idea of *emergent narrative* has become more and more popular [1]. Moreover, there is a growing trend to incorporate cognitive models within the characters involved in virtual stories (e.g., [10], [12]). As a next step in that direction, the current paper aims at building emotion regulation as known from psychology literature into virtual characters. To this end, the informal model by Gross [7] was taken as a basis, and has been formalized using a dynamical system style modeling approach (see also [3] for some initial steps). A virtual environment has been created, which includes a number of virtual agents that have been equipped with the formalized model for emotion regulation. To test the behavior of the model in a prototyping phase, a series of simulation experiments has been performed using the LEADSTO simulation language [2]; in the Vizard Virtual Reality Toolkit [16], such simulations have been visualized in a graphical environment. The resulting movies provide a first indication that the emotion regulation strategies as described by [7] have been implemented successfully within the virtual characters. The simulation results have been compared with the behaviors for different situations as described by Gross [7], [8], and found consistent. Validation involving comparison with detailed empirical data is left for future work.

Concerning related work, an approach in the literature that has similarities to the current approach is [9]. In that paper, a computational model is introduced that can simulate several strategies about how humans cope with emotions, such as ‘positive reinterpretation’ and ‘denial’. Their approach makes use of plan-based causal representations, augmented with decision-theoretic planning techniques, whereas our approach uses dynamical systems representations. Other differences are that they propose a "content model", in which appraisal and regulation operate on rich representations of the emotion-evoking situation, and that their work has been evaluated against clinical data.

Virtual stories involving characters with elaborated cognitive or psychological capabilities can be used for a number of purposes. On the one hand, they may be used for entertainment (e.g., for creating computer games with more complex, unpredictable and more human-like characters). On the other hand, they may be used for educational purposes (e.g., to create a virtual training environment for psychotherapists, which enables them to practice anger management sessions with virtual clients). Further research will investigate whether the model is suitable for such purposes. As soon as these types of challenges will be tackled, also a more precise evaluation will be performed of how humans perceive the current characters (e.g. in terms of believability).

References

1. Aylett, R. (1999). Narrative in Virtual Environments: Towards Emergent Narrative. In: *Working notes of the Narrative Intelligence Symposium*, AAAI Spring Symposium Series. Menlo Park, California, AAAI Press.
2. Bosse, T., Jonker, C.M., Meij, L. van der, and Treur, J. (2007). A Language and Environment for Analysis of Dynamics by SimulaTiOn. *International Journal of AI Tools*. To appear, 2007.
3. Bosse, T., Pontier, M., and Treur, J. (2007). A Dynamical System Modeling Approach to Gross' Model of Emotion Regulation. In: *Proceedings of the 8th International Conference on Cognitive Modeling, ICCM'07*. Taylor and Francis, to appear.
4. Cavazza, M., Charles, F., and Mead, S. (2002). Interacting with virtual characters in interactive storytelling. In: *Proceedings. of the First International Conference on Autonomous Agents and Multi-agent Systems, AAMAS'02*. ACM Press, pp. 318-325.
5. Dautenhahn, K. (1998). Story-Telling in Virtual Environments. Proceedings of the ECAI'98 Workshop on Intelligent Virtual Environments, Brighton, UK.
6. Ekman, P., Friesen, W.V., and Ellsworth, P. (1972). *Emotion in the human face: guidelines for research and integration of Findings*. New York: Pergamon.
7. Gross, J.J. (1998). The Emerging Field of Emotion Regulation: An Integrative Review. *Review of General Psychology*, vol. 2, no. 3, pp. 271-299.
8. Gross, J.J. (2001). Emotion Regulation in Adulthood: Timing is Everything. *Current directions in psychological science*, vol. 10, no. 6, pp. 214-219.
9. Marsella, S., and Gratch, J. (2003). Modeling coping behavior in virtual humans: Don't worry, be happy. In *Proceedings of Second International Joint Conference on Autonomous Agents and Multiagent Systems, AAMAS'03*. ACM Press, pp. 313-320.
10. Marsella, S. C., Johnson, W. L., and LaBore, C. (2000), Interactive Pedagogical Drama. In: *Proceedings of the 4th International Conf. on Autonomous Agents*, ACM Press, pp. 301-308.
11. Ochsner, K.N., and Gross, J.J. (2005). The cognitive control of emotion. *Trends in Cognitive Sciences*, vol. 9, pp. 242-249.
12. Paiva, A., Machado, I., and Prada, R. (2001). Heroes, villains, magicians, ...: Dramatis personae in a virtual story creation environment. In: *Proceedings of the Conference on Intelligent User Interfaces, IUI'01*, pp. 129-136.
13. Theune, M., Faas, S., Heylen D., and Nijholt, A. (2003). The Virtual Storyteller: Story Creation by Intelligent Agents. *Proceedings of Technologies for Interactive Digital Storytelling and Entertainment*, pp. 204-215.

14. Thompson, R.A. (1994). Emotion regulation: A theme in search of definition. In N.A. Fox (ed.), *The development of emotion regulation: Biological and behavioral aspects. Monographs of the Society for Research in Child Development*, vol. 59, pp. 25-52.
15. Van Vugt, H.C., Hoorn, J.F., Konijn, E.A., and De Bie Dimitriadou, A. (2006). Affective affordances: Improving interface character engagement through interaction. *International Journal of Human-Computer Studies,* vol. 64, no. 9, pp. 874-888.
16. Worldviz. *Vizard Virtual Reality Toolkit*. University of California. URL: http://www.worldviz.com/vizard.htm.
17. http://www.cs.vu.nl/~tbosse/virtualstories.

PART V EMBODYING EMOTIONS IN VIRTUAL AGENTS

CHAPTER 10

A Virtual Therapist That Responds Empathically to Your Answers

This chapter appeared as Pontier, M., and Siddiqui, G.F., A Virtual Therapist that responds Empathically to your Answers. In: Prendinger, H. and Ishizuka, M. (eds.), Proceedings of the 8th International Conference on Intelligent Virtual Agents, IVA'08. Lecture Notes in Artificial Intelligence, vol 5208. Springer Verlag, 2008, pp. 417-425.

A Virtual Therapist That Responds Empathically to Your Answers

Matthijs Pontier, Ghazanfar F. Siddiqui

Vrije Universiteit Amsterdam, Department of Artificial Intelligence,
De Boelelaan 1081a, 1081 HV Amsterdam, The Netherlands
Center for Advanced Media Research Amsterdam
Buitenveldertselaan 3, 1082 VA Amsterdam, The Netherlands
{mpontier, ghazanfa}@few.vu.nl
http://www.few.vu.nl/~{mpontier, ghazanfa}

Abstract. Previous research indicates that self-help therapy is an effective method to prevent and treat unipolar depression. While web-based self-help therapy has many advantages, there are also disadvantages to self-help therapy, such as that it misses the possibility to regard the body language of the user, and the lack of personal feedback on the user responses. This study presents a virtual agent that guides the user through the Beck Depression Inventory (BDI) questionnaire, which is used to measure the severity of depression. The agent responds empathically to the answers given by the user, by changing its facial expression. This resembles face to face therapy more than existing web-based self-help therapies. A pilot experiment indicates that the virtual agent has added value for this application.

Keywords: Virtual agent, Self-help therapy, Emotion modeling

1 Introduction

Self-help therapies have been investigated for several decades. Self-help therapy started with bibliotherapy, in which clients follow a therapy from a book. Previous research indicates that this is a very effective form of therapy; e.g., a meta-analysis by Cuijpers [7] concluded that bibliotherapy in unipolar depression is an effective treatment modality, which is no less effective than traditional individual or group therapy.

The advent of new communication technologies, like internet and videoconferencing, can also assist in the field of mental healthcare. Since the last decade, a lot of self-help programs have been delivered through the internet [5], [6], [12]. Several previous studies concluded that self-help therapies are useful and efficient in reducing mental health problems convincingly (e. g., [7], [12]). Compared to traditional therapy methods, web-based self-help may be more efficient and less expensive [4], [9].

Web-based self-help therapy can also be a solution for people who would otherwise not seek help, wishing to avoid the stigma of psychiatric referral or to protect their privacy [13]. The majority of persons with a mental disorder in the general population do not receive any professional mental health services (an estimated 65%) [4]. In many occupations, such as the police force, the fire service and farming, there is much stigma attached to receiving psychological treatment, and the anonymity of web-based self-help therapy would help to overcome this [11]. Also many other people feel a barrier to seek help for their problems through regular health-care systems; e.g., in a study by Spek et al. [12] about internet-based cognitive behavioral therapy for subthreshold depression for people over 50 years old, many participants reported not seeking help through regular health-care systems because they were very concerned about being stigmatized. Patients may be attracted to the idea of working on their own to deal with their problems, thereby avoiding the potential embarrassment of formal psychotherapy [13]. Self-help therapy can also be offered to patients while they are on a waiting list, with the option to receive face to face therapy later, if required [11].

However, there is also critique on internet-based self-help therapy. Drop-out rates from self-help therapy can be high, especially when the use of self-help is unmonitored by a health care practitioner [13]. A wide range of drop-out rates for bibliotherapy have been estimated: from about 7% [7] up to 51.7% [12]. People may miss personal feedback when performing self-help therapy, which might decrease their motivation. By making self-help therapy more similar to face to face therapy, it can become a more personal and entertaining experience, which might decrease drop-out.

Several self-help therapy programs are already available on the internet. Two well-known examples of CBT (Cognitive Behavioral Therapy) programs are ‘BluePages’ and ‘MoodGYM’ [5]. BluePages gives information about the symptoms of depression whereas MoodGYM is designed to prevent depression [5], [6]. However, none of the existing online self-help therapies include a virtual agent that provides a kind of face to face assistance.

There have already been developed several agents in the health-supporting domain. For example, [3] describes a virtual agent that explains health documents to patients.

This study presents an application for performing the Beck Depression Inventory questionnaire [2]. The application is equipped with a virtual agent that responds empathically to the responses of the user. As the virtual agent is emotionally responsive to the answers given by the user throughout the questionnaire, the experience should resemble face to face therapy more than a similar application without a virtual agent.

2 The application

In the application, the user performs the BDI questionnaire [2]. The main goal of the BDI is to measure the characteristic attitudes and symptoms of depression. The BDI is a self-report inventory that consists of 21 multiple-choice questions, and is generally used for measuring the severity of depression. Every question has at least four answer options ranging in intensity from 0 to 3.

The virtual agent asks the questions to the user, and the user selects the appropriate answer from a given drop-down box. This virtual agent has a certain emotional state, consisting of two emotions: happiness and empathy. According to Eisenberg [8], empathy is “*An affective response that stems from the apprehension or comprehension of another’s emotional state or condition, and that is similar to what the other person is feeling or would be expected to feel.*” Because in this application a depression questionnaire is conducted, which means empathy concerns rather sad things, showing empathy consists of showing sadness. If during the questionnaire the user appears to be more depressed the virtual agent will show more sadness, expressed by a relatively sad facial expression. On the other hand, if the user appears to be completely fine, the agent will show a relatively happy facial expression. When the webpage is loaded for the first time, the original emotional state of the virtual agent is loaded, which is a calm emotional state, with very little sadness and an average level of happiness.

2.1. The emotion model of the agent

The virtual agent responds empathically towards the user, by showing the right facial expressions on the answers given by the user. In consultation with clinical psychologists, we defined an impact of these answers as a real number in the domain [-1, 1] on the emotions of the virtual agent, represented by a real number in the domain [0, 1]. Further we defined in consultation with the clinical psychologists how the agent should behave towards the users using these impacts. The impacts are used to detect how the user is feeling. When the user gives a lot of answers that indicate he or she is not feeling well, the agent should show empathy, by showing a sad facial expression, without showing any happiness. When the user is feeling fine, the agent should show a neutral, calm facial expression, with some happiness and no sadness. Because it would be undesirable if the emotions of the agent suddenly shift from very sad to very happy or vice versa, with any change in emotions, the old levels of the emotions are taken into account. If the answers of the user have no impact on the emotions of the agent, its facial expression should slowly return to the original emotional state it had at the start of the application. We have developed the following formula that meets the requirements as described above:

New_emotion = Old_emotion + Decay + Change
Decay = (Original_emotion – Old_emotion) * Decay_factor
Change = ζ * Impact / (1 + (Original_emotion - Old_emotion) * Impact)

New_emotion can be calculated by taking the old emotion, and adding decay and change. Here Old_emotion is the emotion of the virtual agent before the formula is applied. Decay is the size of the decay effect (i.e., how quickly the emotion will move towards the original emotion if the user response has no impact). Change is the change of the emotion of the virtual agent, according to the impact of the answer given by the user.

Decay is calculated by subtracting Old_emotion from Original_emotion, and multiplying the result with the Decay_factor. In this formula, Decay_factor is a variable that determines the size of the decay effect, which is taken 0.1 in this paper. Original_emotion is the emotion of the virtual agent at the start of the application.

Change is calculated by multiplying the impact of the answer given by the user with ζ, and dividing the result by 1 + (Original_emotion - Old_emotion) * Impact. In this formula, ζ is a variable that determines the speed with which the answers given by the user can modify the emotions of the virtual agent. Dividing Impact by 1 + (Original_emotion - Old_emotion) * Impact manages that when the current emotion of the agent deviates more from the original emotion the agent had at the start of the application, and the answer given by the user pushes the emotion of the agent even further away from the original emotion, the change will be relatively smaller as when the user's answer would push the agent's emotion back towards the original emotion with the same impact.

The emotions of the agent during two scenarios are shown in Figure 1. In this figure, along the x-axis the time is given and along the y-axis the levels of emotions are given.

The pink line shows happiness and the blue line shows sadness. In scenario 1, the agent interacts with a severely depressed user, while in scenario 2 the agent interacts with a user who scores average on feelings of depression.

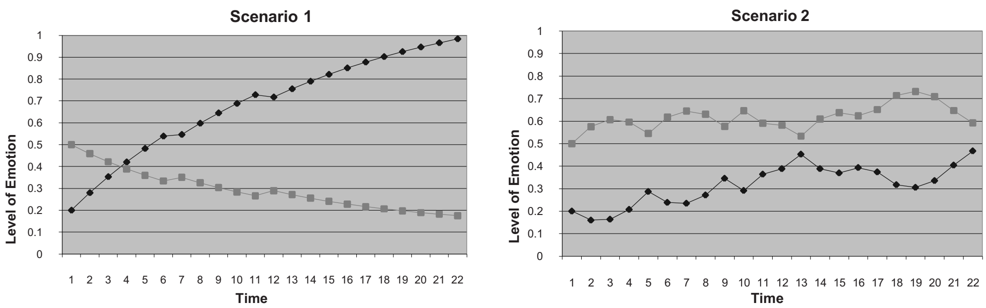


Fig. 1. The emotions of the agent during scenario 1 and scenario 2

As can be seen in Figure 1, initially the virtual agent has a very small level of sadness (0.2) and an average level of happiness (0.5). Each time point the virtual therapist asks a question to the user and gets an answer. This answer affects the emotions of the agent. The user in scenario 1 is severely depressed, coming to a final score of 61 on the BDI questionnaire. During the questionnaire, the agent notices this, which increases her level of sadness and decreases her level of happiness. This results in the agent showing empathy towards the user by means of a sad facial expression.

In scenario 2, the user got an average score on the BDI questionnaire (33). This means the user scores average on feelings of depression. In Figure 1, the different emotional reactions of the agent on the answers of the user in scenario 2 can clearly be seen. On answers that indicate the user is depressed, such as the answer just before time point 5, the agent's level of sadness increases, and the agent's level of happiness decreases. On the other hand, on answers that indicate the user is not depressed, such as the answer just before time point 6, the agent's level of sadness decreases, and the agent's level of happiness increases. At the end of the questionnaire, the agent shows a facial expression with an average level of both sadness and happiness.

2.2 The resulting website

For creating the virtual agent, we used Haptek's peopleputty software [10]. Through this program we created the face of the virtual agent. Further, we created nine different emotional states using the sliders 'happy' and 'sad' given by the software; one for each possible combination of the three levels of the emotions happiness and sadness (showing empathy). We created a webpage for the BDI questionnaire, on which the virtual agent was embedded as a Haptek player. We used JavaScript [1], a scripting language, in combination with scripting commands provided by the Haptek software [10], to control the Haptek player within a web browser.

We performed a pilot experiment to test whether the virtual agent has added value for the application. To recruit participants, we invited people by sending an e-mail with a link to the website. The first page of this website does not show the virtual agent, and contains 8 demographic questions. The next page contains the 21 questions of the BDI questionnaire. The virtual agent is shown on top of this page. Before the first question, the agent introduces itself, and states that it will guide the user through the questionnaire. Instead of showing all the questions in a form, only one question at a time was shown, to let the application more resemble face to face therapy. Each question is shown in a text area below the virtual agent. Below the question, there is a dropdown-box from which the user can select an answer.

Because some of the possible participants were Dutch, we created a Dutch, as well as an English version of the website. In the English version, the virtual agent asks the question using speech. Because at the moment of this study, we did not have a Dutch speech synthesis engine available that included lip-syncing, in the Dutch version the question was only shown in the text area below the virtual agent. We used this shortcoming of the Dutch version to investigate the added value of speech in this application.

Each time the user selected an answer from the drop-down box, the virtual agent changed its emotional state depending on the calculated values of the emotions, as described in section 3.1. Each answer has a score, as described in Section 2, and during the questionnaire these scores are accumulated to calculate the final score. When the user presses the submit button, it proceeds towards the next page, where the virtual agent gives feedback about the final score of the user on the questionnaire, showing an appropriate facial expression. When the final score was below 16, the virtual agent indicates that the user is less depressed as average and shows a facial expression with a low level of sadness and a medium level of happiness. When the user scores between 16-41, the virtual agent indicates that the user scores average on feelings of depression and shows a facial expression with a medium level of sadness and a low level of happiness. When the user scores above 41, the virtual agent indicates that the user scores high on feelings of depression and shows a facial expression with a high level of sadness and a low level of happiness. If the user responded that he or she considers committing suicide, the agent stringently advises the user to contact his or her general practitioner.

After receiving the feedback, the user clicks a button to proceed to the next page, which contains the evaluation form. This page consists of 5 questions about the virtual agent, such as whether the user prefers performing the BDI questionnaire with or without the virtual agent. The virtual agent itself is not shown on this page, to prevent the user from giving socially desirable answers towards the agent.

3 Experiment

We have performed a pilot experiment to test whether the virtual agent has added value for this application.

Participants. The participants were recruited by sending an e-mail with an invitation to participate in the experiment. The participants could choose between a Dutch and an English version of the questionnaire. 28 participants completed the experiment, of which 16 the English, and 12 the Dutch version.

Procedure. First the participants entered some demographical information in a web-form, without the virtual agent. Next the application with the virtual agent was loaded, and the participants performed the BDI questionnaire. When the questionnaire was finished, the participants received feedback from the virtual agent about their result. In the English version, the question was shown in a text area below the agent, and the agent additionally asked the questions to the participants using speech. In the Dutch version however, the virtual agent could not speak, and the text was only shown in the text area below the agent. Finally, the participants filled in an evaluation questionnaire, without the virtual agent present to prevent socially desirable answers towards the agent. The complete procedure can still be performed at [14].

Results. The participants evaluated on an eight-point scale whether they thought the virtual agent was friendly, interested, trustworthy and kind. In both the English and the Dutch version, for all properties, the participants scored the agent just above moderate, as can be seen in Table 1. No statistical differences were found between the English version with voice, and the Dutch version without voice.

Table 1. The score of the virtual agent on several properties on an eight-point scale.

	English		Dutch	
	M	SD	M	SD
Friendly	4.75	1.98	4.91	1.30
Interested	4.50	2.13	4.00	1.95
Trustworthy	4.31	1.66	4.27	1.56
Kind	4.75	1.69	4.55	1.44

Further, the participants answered the question “If you were to administer the same questionnaire, would you rather do this with or without virtual interviewer?” on the evaluation form. For the English version, with speech, 81% of the participants preferred to perform the questionnaire with the virtual agent (sign test, $p = .021$). However, for the Dutch version, without speech, only 64% of the participants preferred to fill in the questionnaire with the virtual agent (sign test, $p = .55$).

The participants were asked to explain their answer in an open question. On this question people gave various responses, but a response that came back several times was that with the virtual agent, it “feels more personal” and that it “feels friendlier”. Participants also indicated that it was more fun to perform the questionnaire with the virtual agent. This indicates that the virtual agent makes it more attractive and entertaining to perform the questionnaire, and adding the virtual agent to a self-help application might decrease drop-out of the self-help therapy.

Participants that preferred to perform the questionnaire without a virtual agent gave as reasons for this that the agent was still too “cold and computer-like”, and with the Dutch version, without speech, that the agent did not have any added value.

The responses on the open question “How do you think the virtual interviewer should be improved?” indicated that there is still a lot of work to do. In the Dutch version, a lot of participants responded that the agent should speak, while in the English version many participants responded that the voice of the agent should be friendlier. In both versions many participants responded the agent should give feedback on each answered question.

4 Discussion

This study presents a virtual agent that guides the user through a questionnaire about depression. The agent responds empathically to the answers given by the user, by changing its facial expression.

A pilot experiment has been performed to test the applicability of a virtual agent in this application. Due to time limitations, the way of recruitment of the participants was not ideal, and the participants are probably not a very good representation of the target group of online self-help applications. When the application has been improved, an extensive validation will need to be performed before it can be used in practice. However, the experiment has led to some interesting results that can be used to determine a direction for further research.

In both the English version, with speech, and the Dutch version, without speech, the participants found the agent moderately friendly, interested, trustworthy and kind. Further, although there were not many participants, an interesting statistical significant result was found. For the English version, the amount of participants who preferred to perform the questionnaire with the virtual agent was significantly bigger than the amount of participants who preferred to perform the questionnaire without the virtual agent. For the Dutch version also more participants preferred to perform the questionnaire with the agent than without, but this result was not significant. However, none of the participants appeared to actually be depressed, and the agent thus will not have shown much obvious empathic expressive behavior. In the Dutch version of the application, without speech, this means the participants just saw a rather passive face above the questions. Given this information, and that there were only 12 participants, it is not very surprising that the for the Dutch questionnaire, the amount of participants that preferred performing the questionnaire with the agent was not significantly bigger than the amount of participants that preferred performing the questionnaire without the agent.

Taking into account many improvements can still be made to the application, the results described above are very promising, and motivate further research in this direction. The response of a participant that he did not feel shy of the virtual interviewer as he would with a real one further indicates nicely the use of this kind of applications. For people who feel uncomfortable with undergoing face-to-face therapy and therefore choose not to seek help, an application like this can be a nice solution.

As also indicated by the responses in the experiment, many improvements can still be made to the application. As pointed out by many participants in the open questions, the agent should provide appropriate feedback after each answer on a question. This should increase the humanness of the agent, enhancing the feelings of a personal, realistic experience during the questionnaire.

Another possible point of improvement is the voice of the agent. The fact that with the Dutch version, without speech, the amount of participants that preferred to perform the questionnaire with the virtual agent was not significantly bigger than the amount of participants that preferred to perform the questionnaire without the virtual agent indicates that this is an important issue. Moreover, in the open questions, many participants gave responses that indicated that speech should be added (in the Dutch version) or improved (in the English version). Since the application should ultimately also be available for Dutch speaking users, possibilities for adding Dutch speech synthesis including lip-syncing should seriously be considered. Also possibilities to create a friendlier voice that is able to show emotions should be considered.

Acknowledgements

We kindly want to thank Annemieke van Straten and Tara Donker for their input to this paper, and Jan Treur, Tibor Bosse and the anonymous reviewers for their comments on earlier drafts of this paper.

References

1. About JavaScript – MDC, http://developer.mozilla.org/en/docs/About_JavaScript
2. Beck, A. T.: Depression: Causes and Treatment. Philadelphia: University of Pennsylvania, Press ISBN 0-8122-1032-8 (1972)
3. Bickmore, T. W., Pfeifer, L. M., and Paasche-Orlow, M. K.: Health Document Explanation by Virtual Agents. Proceedings of the Seventh International Conference on Intelligent Virtual Agents, IVA'07. Lecture Notes in Artificial Intelligence, vol. 4722, pp. 183-196. (2007)
4. Bijl, R. V., and Ravelli, A.: Psychiatric morbidity, service use, and need for care in the general population: results of The Netherlands Mental Health Survey and Incidence Study. American Journal of Public Health, Vol. 90, Iss. 4, pp. 602-607 (2000)
5. Christensen, H., Griffiths, K. M., and Jorm, A. F.: Delivering interventions for depression by using the internet: randomised controlled trial. BMJ, Vol. 328, pp. 265-269 (2004)
6. Christensen, H., Griffiths, K. M., and Korten, A.: Web-based cognitive behavior therapy: analysis of site usage and changes in depression and anxiety scores. Journal of Medical Internet Research, Vol. 4, No. 1: e3. (2002)
7. Cuijpers, P.: Bibliotherapy in unipolar depression, a meta-analysis. Journal of Behavior Therapy & Experimental Psychiatry, Vol. 28, No. 2, pp. 139-147 (1997)
8. Eisenberg, N.: Empathy-related emotional responses, altruism, and their socialization. In: R. J. Davidson & A. Harrington (Eds.). Visions of compassion: Western scientists and Tibetan Buddhists examine human nature, pp. 131-164. London: Oxford University Press (2002)
9. Griffiths, F., Lindenmeyer, A., Powell, J., Lowe, P., and Thorogood, M.: Why Are Health Care Interventions Delivered Over the Internet? A Systematic Review of the Published Literature, Journal of Medical Internet Research, Vol. 8, No. 2: e10. (2006)
10. Haptek, Inc., http://www.haptek.com
11. Peck, D.: Computer-guided cognitive–behavioral therapy for anxiety states. Emerging areas in Anxiety, Vol. 6, No. 4, pp. 166-169 (2007)
12. Spek, V., Cuijpers, P., Nyklíĉek, I., Riper, H., Keyzer, J., and Pop, V.: Internet-based cognitive behavior therapy for emotion and anxiety disorders: a meta-analysis. Psychological Medicine, Vol. 37, pp. 1-10 (2007)
13. Williams, C.: Use of Written Cognitive-Behavioral Therapy Self-Help Materials to treat depression. Advances in Psychiatric Treatment, Vol. 7, pp. 233–240 (2001)
14. http://www.few.vu.nl/~ghazanfa/welcome.php

PART V EMBODYING EMOTIONS IN VIRTUAL AGENTS

CHAPTER 11

An Affective Agent Playing Tic-Tac-Toe as Part of a Healing Environment

This chapter appeared as Pontier, M.A., and Siddiqui, G.F., An Affective Agent Playing Tic-Tac-Toe as Part of a Healing Environment In: J.-J. Yang et al. (eds.), Proceedings of the 12th International Conference on Principles of Practice in Multi-Agent Systems, PRIMA'09, Lecture Notes in Artificial Intelligence, vol 5925. Springer Verlag, 2009, pp. 33-47.

An Affective Agent Playing Tic-Tac-Toe as Part of a Healing Environment

Matthijs Pontier[1, 2] and Ghazanfar Farooq Siddiqui[1, 2, 3]

[1] VU University Amsterdam, Center for Advanced Media Research Amsterdam, De Boelelaan 1083, 1081HV Amsterdam, The Netherlands
[2] VU University Amsterdam, Department of Artificial Intelligence, De Boelelaan 1083, 1081HV Amsterdam, The Netherlands
[3] Quaid-i-Azam University Islamabad, 45320, Pakistan.
{mpr210, ghazanfa}@few.vu.nl
http://www.few.vu.nl/~{mpr210, ghazanfa}

Abstract. There is a growing belief that the environment plays an important role in the healing process of patients, supported by empirical findings. Previous research showed that psychological stress caused by loneliness can be reduced by artificial companions. As a pilot application for this purpose, this paper presents an affective agent playing tic-tac-toe with the user. Experimenting with a number of agents under different parameter settings shows the agent is able to show human-like emotional behavior, and can make decisions based on rationality as well as on affective influences. After discussing the application with clinical experts and making improvements where needed, the application can be tested in a clinical setting in future research.

Keywords: Cognitive Modeling, Emotion Modeling, Healing Environment

1 Introduction

Many people do not like the atmosphere in hospitals. Since two decades, there is a growing belief that not only the health care itself, but also the environment affects the healing process of the patients. This has increased the interest in healing environments. The role of the environment in the healing process is a growing concern among health care providers, environmental psychologists, consultants and architects. Among them the consensus is growing that not only the level of care, but also the design of the health care facility affects the wellness of its patients [7].

Researchers are finding that making changes and additions to the physical and social environment of the health care facility, thereby taking the patient into account, can positively influence patients' outcomes (e.g., [2], [6], [18], [22], [25]). Moreover, health care professionals are finding that changes in design can enhance recovery in patients, and reduce the length of their stay in the hospital [15]. On the other hand, researchers are also finding that unfamiliar environments in clinics, hospitals, and nursing homes can produce psychological stress that can negatively affect healing and wellness. Poor design has even been linked to negative effects on the patient, such as anxiety, delirium, elevated blood pressure levels, and an increased intake of pain drugs [23].

One factor that can be reduced by a healing environment is psychosocial stress. An important predictor of psychosocial stress is loneliness [12]. Loneliness is a common problem frequently encountered in the elderly in long-term care facilities. Many people that are staying in a long-term care facility lack social interaction. Artificial toys can be used to reduce loneliness. Previous research showed that animal-shaped toys can be useful as a tool for occupational therapy (e.g., [18], [26], [27]). Robot animal therapy has been widely investigated. For example, Dautenhahn and Robins [20], [28] used mobile robots and dolls respectively to treat autistic children. Wada and Shibata developed Paro [27], a robot shaped like a baby-seal that interacts with users to encourage positive mental effects. Interaction with Paro has been shown to improve users' moods, making them more active and communicative with each other and caregivers. Research groups have used Paro for therapy at eldercare facilities and with those having Alzheimer's disease [14], [17]. Banks et al. [2] showed that animal-assisted therapy with an AIBO dog helped just as good for reducing loneliness as therapy with a living dog. In their paper they indicate that AIBO was not used to its full capacity and that if more options were used, its effects might be further enhanced.

Over the past decade, a lot of novel work on computational models of emotion in virtual agents can be observed. Nevertheless, compared to human affective complexity, current emotion models of virtual agents are still quite simple. If an artificial companion demonstrates human-like emotional behavior, this might increase its ability to reduce loneliness of patients in a long-term care facility, as part of a healing environment.
In our paper, we present a virtual agent that could be seen as a pilot application for this purpose. The artificial companion is an affective virtual agent that can play tic-tac-toe, equipped with Silicon Coppélia [19], an integration of three affect-related models as proposed in [3]. Because it is equipped with these affect-related models, it can show human-like emotional behavior. Therefore, it might be a useful to serve as an artificial interaction partner for patients in a long-term care facility.

2 The Application

The application presented in this paper is an affective virtual agent that can play tic-tac-toe against the user. The object of tic-tac-toe is to get three in a row on a three by three game board. You play on a three by three game board. Players alternate placing X's and O's on the game board until one of the players has three in a row, or all nine squares on the board are filled, which means the game ends in a tie. For creating the virtual agent, we used Haptek's peopleputty software [11]. Through this program we created the face of the virtual agent. The agent simulates 5 emotions: joy, distress, hope, fear and surprise, which can be expressed with either a low or a high intensity.

We created 32 (2^5) different emotional states using peopleputty; one for each possible combination of two levels of intensity of the five emotions simulated by the agent. We created a webpage for the application, on which the virtual agent was embedded as a Haptek player. We used JavaScript [1], a scripting language, in combination with scripting commands provided by the Haptek software [11], to control the Haptek player within a web browser.

Figure 1 shows the resulting website. In Figure 1, the agent, playing O's, just lost a game, and therefore looks sad. The website shows in the top left the agent which the user plays against. If the agent speaks a message, this is additionally shown in a textbox that is shown below the agent. Below this textbox, the ambition levels of the agent for winning and losing, and the importance of the current game for the agent are shown.

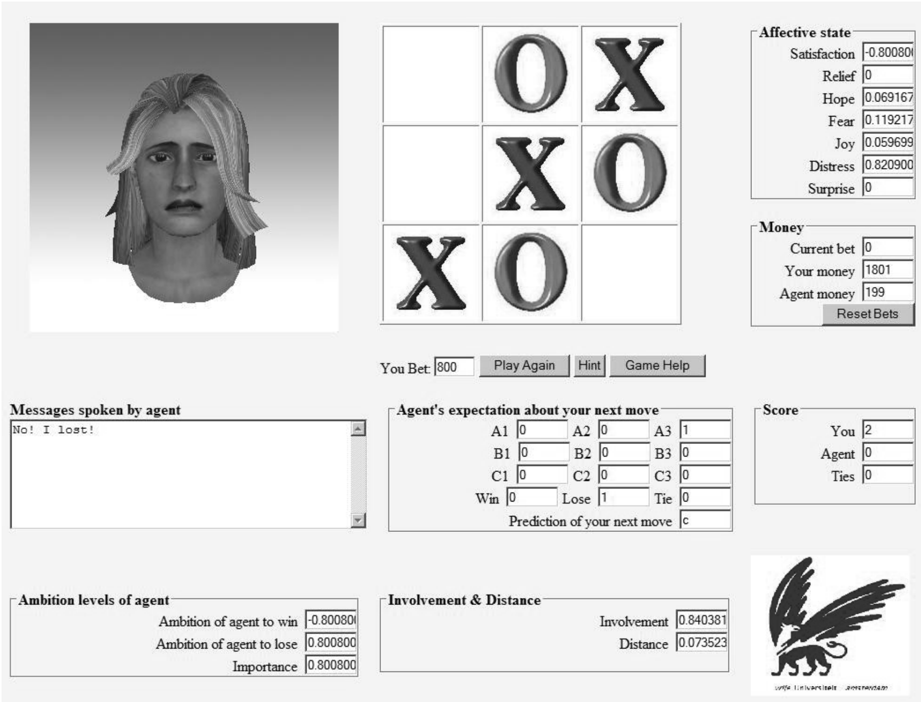


Fig. 1. The website with the tic-tac-toe application.

Right next to this, the level of involvement and distance from the agent towards the user are shown. Right next to the textbox, the predictions of the agent about the expected next move and the outcome of the game are shown. Above this, just below the tic-tac-toe board, the user can enter its bet, and there is a ‘play again’ button which the user can click on to play a new game with the inserted amount of money as bet. Additionally, there is a hint button, which makes the agent give a hint to the user about in which square to make a move. Further, there is a ‘game help’ button. If the user clicks this button, the agent will explain the rules of the game.

Because the purpose of this paper is to show how the application works, the affective state is not only shown by means of a facial expression of the agent, but the emotion variables are also shown numerically on the top right. If the application would be designed to be used by human users, these numerical values would not be shown, and only the facial expression would be visible to the user. Below the affective state in numbers, the amount of money that is currently played for (current bet), the amount of money of the user (your money), and the amount of money of the agent (agent money) are shown. There is also a ‘reset bets’ button, which resets the bets to the starting values. Below this, the number of games won by the user, the number of games won by the agent, and the number of ties is shown.

The tic-tac-toe board is on the top center of the website, right next to the agent. The user can make a move by clicking on one of the squares, on which the agent will react by performing its own move. After each move of the human user, the agent speaks a message, which is additionally displayed in the text-area below its face, depending on the emotional state of the agent. If the game has finished, the amount of money bet for will be added to the winner, and subtracted from the loser. The agent speaks a message, depending on the outcome of the game, and its emotional state. The user can enter a new bet and click the 'play again' button to play another game.

2.1 The models incorporated in the agent

This virtual agent presented in this paper was constructed by incorporating Silicon Coppélia [19], an integration of three affect-related models into an existing virtual agent that can play tic-tac-toe [29]. The three models that were integrated into Silicon Coppélia as suggested in [3] were:

1. EMA [9], [16], a model to create agents that exhibit and cope with (negative) affect based on Smith & Lazarus' theory of emotion [21]. A graphical representation of EMA is shown in Figure 2.

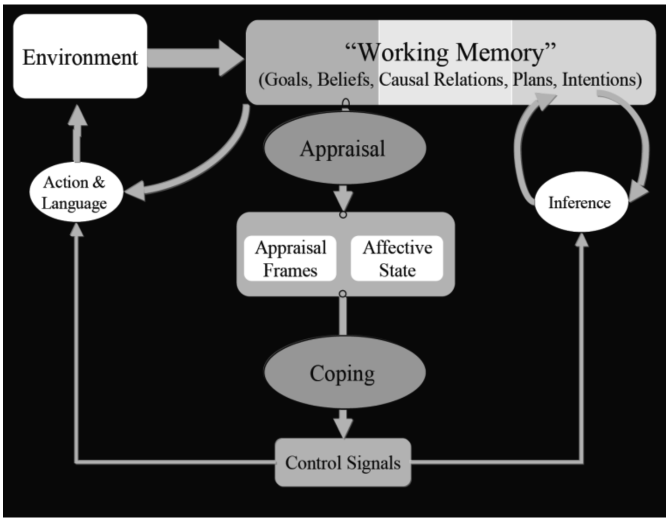


Fig. 2. A graphical representation of EMA

2. CoMERG [4] (the Cognitive Model for Emotion Regulation based on Gross), which can simulate different emotion regulation strategies explained by Gross [10] using a set of logical rules and difference equations. Figure 3 shows a graphical representation of the emotion regulation model by Gross.

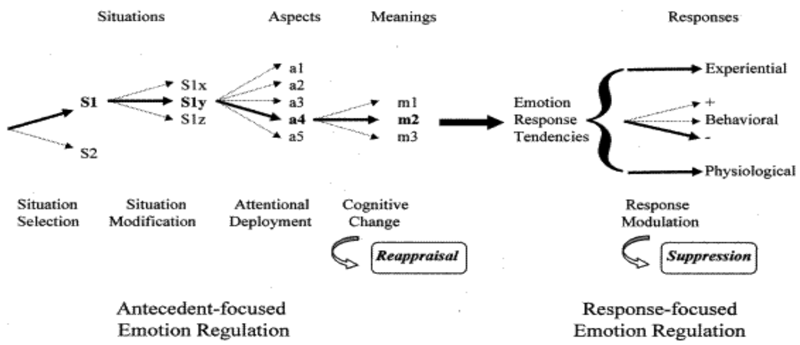


Fig. 3. A graphical representation of the emotion regulation model by Gross

3. I-PEFiC^{ADM} [13], a model for building agents that can trade rational for affective choices based on the concern-driven theory of Frijda [8]. A graphical representation of I-PEFiC^{ADM} is shown in Figure 4.

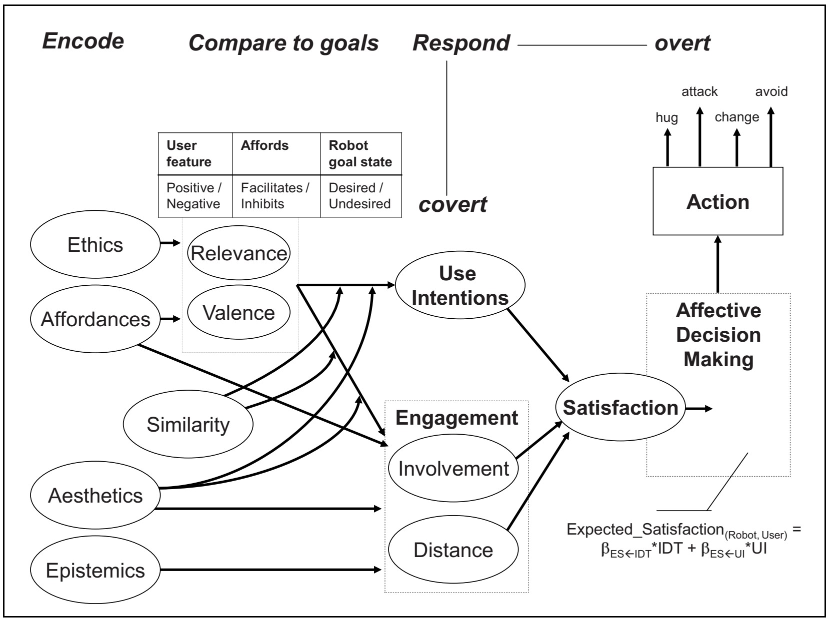


Fig. 4. A graphical representation of I-PEFiC^{ADM}. Curved arrows indicate interaction effects.

Integrating these models enabled agents in simulation experiments to show richer interaction than they could with any of the models alone. Using the combined model, they could simulate emotions based on beliefs about states in the world, and how these states affect their goals. The model was also used to simulate affective decision making processes, in which decisions are made not only based on rationality, but also on affective influences, enabling the agent to make irrational decisions where appropriate. Further, emotion regulation strategies can be applied by the agent, to regulate its (simulated) emotions [19].

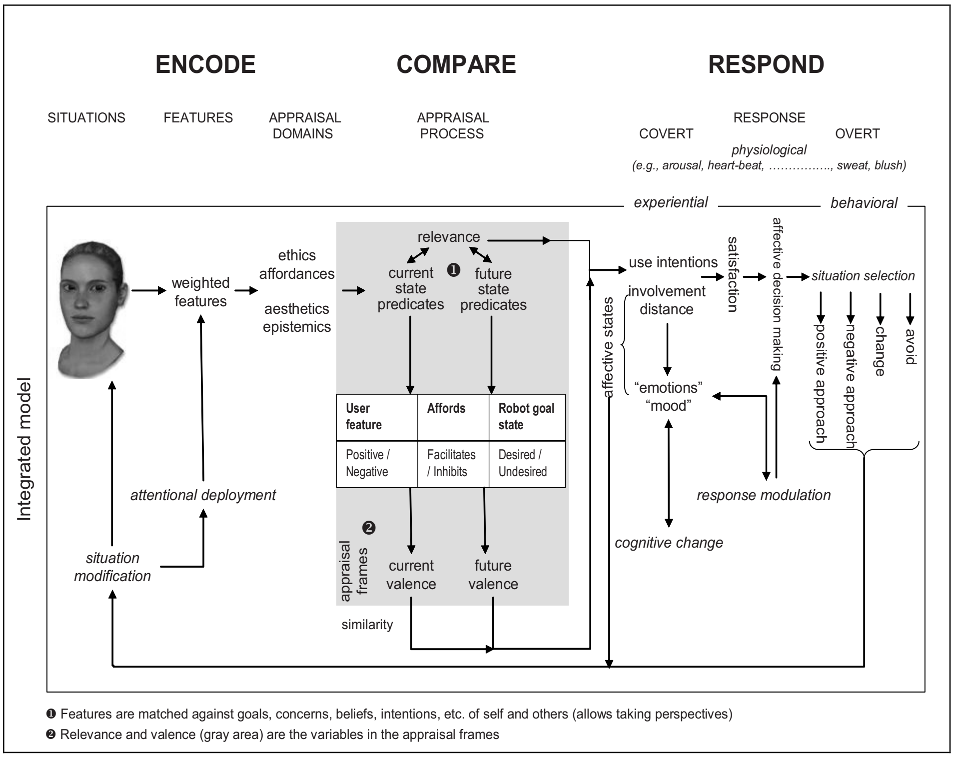


Fig. 5. A graphical representation of Silicon Coppélia.

Figure 5 shows a graphical representation of Silicon Coppélia. On the far left of this figure, we see a virtual agent. The agent develops state predicates about her opponent. The agent acquires personal meaning for her opponent because she compares the features of her opponent with her own personal goals, beliefs, and concerns. This establishes her relevance and valence to her opponent. While relevance determines the intensity of affect, valence governs its direction. The agent can also take perspectives and look at others through the eyes of another person or agent.

When the initial appraisal process is completed, the agent is ready to affectively respond to her opponent. Relevance and valence form an appraisal frame that feeds into her involvement and distance towards her opponent. Inside, the agent will 'experience' several (perhaps ambiguous) emotions.

During affective decision making, the agent selects the move in the game that promises the highest expected satisfaction. The performed action leads to a new situation, and after her opponent also made his/her move, the model loops until the game has finished.

2.2 Determining which action to take

In Silicon Coppélia [19], the agents perceive each other's features according to I-PEFiC [24], by multiplying a designed value (a value the designer expects to raise in another agent), and a bias for perceiving this feature. For this application, as starting values, all biases were set to the neutral value of 1 for the agent. The designed values of the human user were set as can be found in Table 1. These values were chosen arbitrarily, and in an application with real users they should be reconsidered together with an expert. This will lead the agent to perceive the human user with the following variables:

Table 1. Designed values for perceiving the features of the human user.

Feature	Designed value of human user
Good	0.6
Bad	0.2
Beautiful	0.5
Ugly	0.3
Realistic	0.9
Unrealistic	0.1
Intended Aid	0.7
Intended Obstacle	0.4

In Silicon Coppélia, the agents also have beliefs that features of other agents affect certain goal-states in the world. For this application, the possible goal-states are 'agent wins' and 'human wins'. The beliefs about features facilitating these goal-states were set as can be seen in Table 2:

Table 2. Designed values for perceiving the features of the human user.

Feature or Action	Facilitates 'agent wins'	Facilitates 'human wins'
Good	0.5	-0.5
Bad	-0.5	0.5
Beautiful	0.1	0.1
Ugly	0.1	0.1
Realistic	0.1	0.1
Unrealistic	0.1	0.1
Intended Aid	0.7	-0.7
Intended Obstacle	-0.7	0.7

The agent also has beliefs about actions facilitating goal-states. As there are nine squares on the board, there are nine possible actions for the agent: putting an O in each possible place. In Silicon Coppélia, the agents have beliefs that action facilitate goal-states in the range [-1, 1], where -1 means the agent beliefs the action strongly inhibits the goal-state, and 1 means the agent beliefs the action strongly facilitates the goal-state. Each turn, the beliefs that actions facilitate goal-states are calculated based on a heuristic that estimates the chances of winning when a certain move is made. If the agent can make three O's in a row, the belief that performing this action facilitates 'agent wins' is set to 1, and the belief that this facilitates 'human wins' is set to -1. If the human user has two X's in a row, and the third square in this row is still empty, the agent will belief that putting an O on this square facilitates 'agent wins' with a value of 0.9, and inhibits 'human wins' with a value of -0.9. If a square is already occupied, the belief that putting an O on this square facilitates a goal is set to 0 for both goals, as it is an illegal move. If none of these rules apply for a square, the belief that putting an O on the middle square (B2) facilitates 'agent wins' is set to 0.5, and that it facilitates 'human wins' to -0.5. Similarly, the belief that putting an O on corner squares (A1, A3, C1 and C3) facilitates 'agent wins' is set to 0.3 and for 'human wins' this is set to -0.3. Finally,

the belief that putting an O on the remaining squares (A2, B1, B3 and C2) facilitates 'agent wins' is set to 0.1 and for 'human wins' this is set to -0.1. For actions which facilitate desired goals, and inhibit undesired goals (i.e., actions with a high expected utility, generalized over multiple goals), strong action tendencies are calculated [19].
In Silicon Coppélia, each action has a level of positivity and a level of negativity. In this application, the level of negativity of each action is defined as the belief of the agent that the action facilitates winning the game. The level of negativity of the action is defined as 1 minus this belief.

Using these variables, the affective decision-making of Silicon Coppélia [19] is used to determine the action of the agent. The expected satisfaction is calculated using the following formula, and the action with the highest level of expected satisfaction is picked.

$$\begin{aligned}&\text{ExpectedSatisfaction(Action)} =\\&\quad w_{at} * \text{Action\_Tendency} +\\&\quad w_{pos} * (1 - \text{abs}(\text{positivity} - \text{bias}_I * \text{Involvement})) +\\&\quad w_{neg} * (1 - \text{abs}(\text{negativity} - \text{bias}_D * \text{Distance}))\end{aligned}$$

The agent will search for the action with the level of positivity that is closest to the level of (biased) involvement towards the user, the level of negativity closest to the level of (biased) distance towards the user, and the strongest action tendency. The importance of positivity, negativity and action tendency for selecting an action can be adjusted by changing the weights (w_{pos}, w_{neg}, and w_{at} respectively). If an agent wants to perform more positive actions, it can, for example, increase its bias for involvement, and decrease its bias for distance. This way, the agent will prefer more positive and less negative actions.

Note that this way, the agent can also deliberately lose by setting a high ambition level for the goal 'human wins', or because the agent is too involved with the user to try to win each game. The agent can also determine its ambition level for winning on the outcomes of the previous games, and the amount of money that is played for. For example, it can try to win the game if it has less money than the human user, and deliberately try to lose if it has more money than the human user. The agent can also determine the importance of each game, by dividing the amount that is played for by the total amount of money the agent has left. This importance can then be the deviation from 0 in the ambition level.

2.3 Calculating the emotions of the agent

The agent simulates some emotions while playing the game, based on the actions that are being performed by the user, the perceived likelihood of winning and losing, and the outcome of the game. Hope and fear are calculated each time the agent has made its move and the human user is on turn. The hope and fear of the agent are based on the perceived likelihood it will win or lose the game. If the human user is on turn and can make a winning move, it estimates the likelihood for losing 0.8, the likelihood for a tie 0.1, and the likelihood for winning 0.1. If the agent could make a winning move if it would be on turn (but it cannot, because the human user is), it will estimate its likelihood for winning 0.5, the likelihood for a tie 0.4 and its likelihood for losing 0.1. Otherwise, the likelihood for winning and losing are both estimated 0.3 by the agent and the likelihood for a tie is estimated 0.4.

The found likelihoods are used in the following function to calculate the hope for a goal. This function is similar to the function described in [5].

IF f >= likelihood THEN hope_for_goal =
-0,25 * (cos(1 / f * π * likelihood(goal)) -1,5) * ambition_level(goal)

IF f < likelihood THEN hope_for_goal =
-0.25 * (cos(1 / (1-f) * π * (1-likelihood(goal))) -1.5) * ambition_level(goal)

These functions differ from most approaches present in the literature, since their top is not situated at the point where the likelihood is 0. In these functions, f is a shaping parameter (in the domain [0, 1]) that can be used to manipulate the location of the top of the hope curve. The value of this parameter may differ per individual, and represents 'fatalism' (or 'pessimism'): the top of the likelihood/hope-curve is always situated at the point where likelihood = f. Thus, for an f close to 1, the top of the curve is situated to the extreme right (representing persons that only 'dare' to hope for events with high probabilities). Similarly, for an f close to 0, the top of the curve is situated to the extreme left (representing persons that already start hoping for events with low probabilities). In this paper, f is set at 0.5. We chose a smooth function instead of a linear function, because this function best matches the emotion curves found in humans. Furthermore, a higher ambition level simply leads to a higher hope (which is standard in the literature). If the ambition level is negative (i.e., the goal is undesired), the outcome of hope_for_goal will be a negative value.

The following algorithm is performed to the found values for hope_for_goal

1. Sort the values in two lists: [0→1] and [0→-1]
2. Start with 0 and take the mean of the value you have and the next value in the list. Continue until the list is finished. Do this for both the negative and the positive list.
3. Hope = Outcome positive list.   Fear = abs(Outcome negative list).

The values are sorted in a list with positive hope_for_goal's (i.e., hope for desired goals), and negative hope_for_goal's (i.e., fear for undesired goals). For both the lists, 0 is the starting point, and the mean of the value you have and the next value in the list (where the next value is the value closest to 0 that is left in the list) is picked until the end of the list is reached. The new level of hope for the agent is the outcome of the positive list, and the new level of fear for the agent is the absolute value of the outcome of the negative list.

The joy and distress of the agent are based on reaching or not reaching desired or undesired goal-states. If a goal-state becomes true (i.e., the agent wins or the human user wins), the levels of joy and distress are calculated by performing the following formulas:

```
IF ambition_level(goal) >= 0 THEN:
   new_joy        = old_joy + mf_joy * ambition_level(goal) * (1-old_joy)
   new_distress   = old_distress + mf_distress * -ambition_level(goal) * old_distress

IF ambition_level(goal) < 0 THEN:
   new_joy        = old_joy + mf_joy * ambition_level(goal) * old_joy
   new_distress   = old_distress + mf_distress * -ambition_level(goal) * (1-old_distress)
```

In this formula, mf_joy and mf_distress are modification factors that determine how quickly joy and distress change if the agent wins or loses the game. In this paper, the values were both set at 1. These modification factors are multiplied with the impact value, which is ambition_level(goal) for joy and -ambition_level(goal) for distress. This way, if a desired goal is reached, this will increase joy and decrease distress, and reaching an undesired goal will decrease joy and increase distress. Multiplying with limiter (1-old_joy) for joy and old_distress for distress if the goal is desired manages the formula does not go out of range. Further, it manages that if an agent's level of joy or distress approaches an extreme value, it will be harder to push it further to the extreme, and easier to get it back to a less extreme value. If the reached goal is undesired, old_joy is used as limiter for joy and (1-old_distress) as limiter for distress, because the values of joy and distress will move in the opposite direction as when the goal is desired.

The level of surprise is calculated in a similar manner as in [5], [29]. To calculate the level of surprise during the game, the agent generates expectations about which move the user will make. If a square is free, that square gets a point. If the user can make three in a row, or prevent the agent from making three in a row the next turn, the square of that move gets 1 extra point. If the user has one X in a row, and the remaining squares of that row are free, those squares get 0.5 point. After all the squares have got their points, the sum of all points of the squares is normalized to 1. The resulting values for each square are the predicted likelihoods of the human making a move on that square. If the user makes a move on a certain square, the level of surprise for the agent is 1 – likelihood(move). If the game finishes, the level of surprise for the agent is 1 minus the perceived likelihood that the game would end that way.

After each move of the human user, the agent speaks a message, depending on the level of surprise. In the system, there is a small database of messages, labeled with certain emotion intensities. If the level of surprise is very low, it will show that it expected this move, with a message like 'I thought you would do that' or 'A predictable move'. On the other hand, if the agent is very surprised by the move of the user, it will speak a more surprised message, like 'That move surprised me!' or 'You are full of surprises.'

All five emotions inserted in the system (joy, distress, hope, fear and surprise) are simulated in parallel. If the level of joy, distress or surprise is below 0.5, a low intensity of the emotion is shown by the agent. If the level of joy, distress, or surprise is greater or equal than 0.5, a high intensity of the emotion is shown by the agent. Because playing the tic-tac-toe game rarely leads to extreme values of hope and fear in the agent, for hope and fear this boundary is set to 0.25.

After each game has ended, the level of satisfaction for the agent is calculated in the range [-1, 1]. If the agent wins, the level of satisfaction will be the ambition level for 'agent wins'. Similarly, if the human user wins, the level of satisfaction will be the ambition level for 'human wins'. If the game ends in a tie, the importance of the game is calculated by dividing the amount that is played for by the total amount of money of the agent. The satisfaction of the agent after a tied game is then calculated by multiplying the importance of the game with 0.5.

Also a level of relief is calculated for the agent after each game, in the range [-1, 1], by multiplying the level of satisfaction with the level of surprise. Further, a message is displayed, based on the outcome of the game and the level of satisfaction and relief, the agent speaks a message in a similar way as after each move. If the agent wins, and the level of relief is low, and the level of satisfaction is low, it will display a neutral message like 'I win'. If the level of satisfaction is higher, it will display a more enthusiastic message like 'Superb match for me!' If the level of relief is higher, it speaks this relief with a message like 'I won, that is such a huge relief'. If the agent loses, and has a relatively neutral level of satisfaction close to 0, it will display a message like ‘You won’. If the agent is very dissatisfied, with a value close to -1, it will speak a more dramatic message, like ‘No! I lost everything!’ If the game ends in a tie, the agent speaks a neutral message like ‘We tied’. If the agent’s ambition was to win, and it wins, it will show a happy facial expression. If it loses, it will look sad. If the agent’s ambition was to lose, and it wins, it shows a sad facial expression. If its ambition was to lose and it loses, the facial expression will be happy.

3 Results

To test the application, the behavior of the agent has been tested under various parameter settings. All agents experimented with can be found at [30].

Agent 1: The agent tries to win

The ambition level for winning of agent 1 is set to 1, and its ambition level for losing is -1. The weight of the affective influences in the decision making process is set to 0. Under these parameter settings, the agent will always try to win. Because in tic-tac-toe it is impossible to lose if you play it right and you do not want to lose, it is impossible to win of the agent. The best game outcome that can be achieved is a draw.

If the agent wins, it will increase its joy and decrease its distress. If the agent loses, it will decrease its joy and increase its distress. If you make a move the agent does not expect, or the game ends otherwise than expected, it will be surprised. The expectations of the agent can be seen on the website. If the agent thinks it is likely that it will win, it will have a relatively high level of hope, and if it thinks it is likely that it will lose, it will have a relatively high level of fear.

Agent 2: The agent deliberately tries to lose

Agent 2 has an ambition level for winning of -1, and an ambition level for losing of 1. This means for the agent, winning is an undesired goal, and losing is a desired goal. The weight of the affective influences in the decision making process have set to 0, so the agent will always try to lose. The only way to let the agent win is to make sure you don’t make three in a row, and with the last move of the agent, it can do nothing else than make the winning move. Because the agent wants to lose, it will increase its level of joy and decrease its level of distress when it loses. If it wins, it will decrease its level of joy, and increase its level of distress.

Agent 3: The agent decides whether it wants to win based on its money

For agent 3, the ambition level for winning and losing is dependent on its amount of money compared to the amount of money of the human. If the agent has more money than the human, it will have an ambition level for winning of -1*importance and an ambition level for losing of 1*importance. If the amount of money of the agent is less or equal than that of the human, the agent will have an ambition level for winning of

1*importance and an ambition level for losing of -1*importance. The weight of the affective influences of the agent is set to 0. This causes the agent to try to win, unless it already has more money than the human user. If the agent has the ambition to win and it does, it will decrease its joy, and decrease its distress, and if it loses it will decrease its joy and increase its distress. However, if the agent has the ambition to lose, and it wins, it will decrease its joy and increase its distress, and vice versa if it loses. How big the increases and decreases of joy and distress are, depends on the importance of the game.

Agent 4: The agent is too involved with the user to win

Agent 4 is very involved with its user. It is programmed to perceive its user as good, beautiful, realistic, and intending to aid (designed values set to 1). It is also programmed to perceive its user as not bad, not ugly, not unrealistic, and not intending to obstruct. This causes the agent to be very involved with the user with a value of 0.85, and not much at a distance towards the user, with a value of 0.08 at the start of the simulation. The ambition level of the agent to win is defined as the importance of the game, and the ambition level to lose as the negation of this importance. The weight of rational influences in the decision-making process is set to 0. This makes the agent want to perform actions towards the user with a high level of positivity, and a low level of negativity. Because actions to win the game have a relatively low level of positivity and a high level of negativity, the agent will perform actions that facilitate losing the game. Because this agent always has a positive ambition to win and a negative ambition to lose, winning will always increase its level of joy, and decrease its level of distress, and losing will always decrease its level of joy, and increase its level of distress. How big this increase or decrease is, is dependent on the importance of the game.

Agent 5: The agent is balanced, and wins sometimes, and loses sometimes

Agent 5 perceives the user with values as can be seen in Table 1. This leads the agent to be involved with the user with a value of 0.63, and to be at a distance with the user with a value of 0.26 at the start of the simulation. The ambition level of the agent to win is defined as the importance of the game, and the ambition level to lose as the negation of this importance. The rational influences in the decision-making process are set to 0.8, and the influences of positivity and negativity of action are both set to 0.1. Despite having a higher ambition to win than to lose under all circumstances, the agent will be too involved with the user to try to win in a game for a small amount of money. However, if the agent's money is almost gone, or the game is about a lot of money, the agent finds the game so important that it will do its best to try to win.

Similarly to agent 4, if agent 5 wins, this will increase joy and decrease distress, and vice versa if it loses.

4 Discussion

This study presents an affective virtual agent that can play tic-tac-toe. Because it is equipped with Silicon Coppélia [19], an integration of three affect-related models as suggested in [3], it can show human-like emotional behavior.

We created five different agents, each with different parameter settings, to test the behavior under various conditions. We manipulated the ambition levels of the agent, and thereby created agent 1, an agent that always tries to win, and agent 2, an agent that always deliberately tries to lose. Agent 3 determines its ambition level for winning and losing on whether it has more money than the user or not. If the user has more money than the agent, it will deliberately try to lose, but otherwise it will try to win. Agent 4 bases its decisions in the game on emotions, and because it is designed to be very involved with the user, it will perform actions that facilitate the user winning the game. Agent 5 bases its decisions partly on emotions, and partly on rationality. Agent 5 always has more ambition to win than to lose. How big this difference in ambition is, is dependent on the amount of money that is played for. This results in the agent trying to win if the agent plays for a big amount of money, or when its money is almost gone. However, if the game is only about a small amount of money, the agent will be too involved with the user to try to win. Based on whether the agent reaches its goals (winning and losing when the agent has ambitions to win or lose), the likelihood of these goals, and the expectedness of the move of the user and the outcome of a game, the emotions joy, distress, hope, fear and surprise are simulated and shown by the agent by means of facial expressions.

This virtual agent presented in this paper should be seen as a pilot application. Many improvements can still be made, such as giving feedback in a more sophisticated manner. Before it can be tested in a clinical setting, we should first discuss with experts where the application could be improved. We should also discuss with them what type of behavior the agent should show under which conditions, and adjust the parameter settings to meet these requirements. After that, user studies should indicate under which parameter settings participants find the agent most human-like.

However, experimenting with a number of agents, each with different parameter settings indicates that a realistic affective agent playing tic-tac-toe can be created.
Previous research already showed that interacting with a robot pet could decrease loneliness in patients staying in a long-term care facility. An artificial interaction partner that can show human-like emotional behavior might even have a greater beneficial effect on decreasing loneliness in patients. In future research, we intend to perform user studies to show whether this really is the case.

Acknowledgments

We would like to thank Tibor Bosse and Edwin Zwanenburg for their input to the paper, and Edwin Zwanenburg for implementing the game environment.

References

1. About JavaScript – MDC, http://developer.mozilla.org/en/docs/About_JavaScript
2. Banks, M.R., Willoughby, L.M., Banks, W.A.: Animal-Assisted Therapy and Loneliness in Nursing Homes: Use of Robotic versus Living Dogs. Journal of the American Medical Directors Association, vol. 9, 173--177 (2008)
3. Bosse, T., Gratch, J., Hoorn, J.F., Pontier, M.A., and Siddiqui, G.F. (submitted).: Coppélius' concoction: Similarity and complementarity among three affect-related agent models. International Conference on Agents and Artificial Intelligence, ICAART 2010 (2010)
4. Bosse, T., Pontier, M.A., Treur, J.: A Dynamical System Modeling Approach to Gross´ Model of Emotion Regulation. In: Lewis, R.L., Polk, T.A., and Laird, J.E. (eds.), Proceedings of the 8th International Conference on Cognitive Modeling, ICCM'07, 187--192, Taylor and Francis (2007)
5. Bosse, T., Zwanenburg, E.: There's Always Hope: Enhancing Agent Believability through Expectation-Based Emotions. In: Pantic, M., Nijholt, A., Cohn, J. (eds.), Proceedings of the 2009 International Conference on Affective Computing and Intelligent Interaction, ACII'09, 111—118, IEEE Computer Society Press (2009)
6. Davidson, A.W.: Banking on the Environment to Promote Human Well-being. In: Seidel, A.D. (Ed.), Banking on design? Proceedings of the 25th annual conference of the Environmental Design Research Association, 62--66. Oklahoma City, OK: EDRA. (1994)
7. Devlin, A.S., Arneill, A.B.: Health Care Environments and Patient Outcomes: A Review of the Literature. Environment and Behavior, vol. 35, no. 3, 665--694 (2003)
8. Frijda, N.H., The Emotions. Cambridge University, New York (1986)
9. Gratch, J., Marsella, S.: Evaluating a computational model of emotion. Journal of Autonomous Agents and Multiagent Systems (Special issue on the best of AAMAS 2004), vol. 11, no. 1, 23--43 (2006)
10. Gross, J.J.: Emotion Regulation in Adulthood: Timing is Everything. Current directions in psychological science, Vol. 10, No. 6, 214--219 (2001)
11. Haptek, Inc., http://www.haptek.com
12. Hawkli, L.C., Masi, C.M., Berry, J.D., Cacioppo, J.T.: Loneliness Is a Unique Predictor of Age-Related Differences in Systolic Blood Pressure. Psychology and Aging, vol. 21, no. 1, 152--164. DOI: 10.1037/0882-7974.21.1.152 (2006)
13. Hoorn, J.F., Pontier, M.A., Siddiqui, G.F.: When the user is instrumental to robot goals. First try: Agent uses agent. Proceedings of IEEE/WIC/ACM Web Intelligence and Intelligent Agent Technology 2008, WI-IAT '08, IEEE/WIC/ACM, Sydney AU, 296--301. ISBN: 978-0-7695-3496-1. DOI: 10.1109/WIIAT.2008.113 (2008)
14. Kidd, C., Taggart, W., and Turkle, S.: A Social Robot to Encourage Social Interaction among the Elderly. Proceedings of IEEE ICRA. pp. 3972—3976 (2006)
15. Lemprecht, B.: The gap between design and healing. Metropolis, vol. 77, 123 (1996)
16. Marsella, S., Gratch, J.: EMA: A Model of Emotional Dynamics. Cognitive Systems Research, vol. 10, no. 1, 70--90 (2009)
17. Marti, P. Bacigalupo, M., Giusti, L., Mennecozzi, C.: Socially Assistive Robotics in the Treatment of Behavioral and Psychological Symptoms of Dementia. Proceedings of BioRob, 438--488 (2006)
18. Nakajima, K., Nakamura, K., Yonemitsu, S., Oikawa, D., Ito, A., Higashi, Y., Fujimoto, T., Nambu, A., Tamura, T.: Animal-shaped toys as therapeutic tools for patients with severe dementia. Engineering in Medicine and Biology Society, 2001: Proceedings of the 23rd Annual International Conference of the IEEE, vol. 4, 3796--3798 (2001)
19. Pontier, M.A., Siddiqui, G.F.: Silicon Coppélia: Integrating three affect-related models for establishing richer agent interaction. Proceedings of the International Conference on Intelligent Agent Technology, IAT'09. To appear. (2009)

20. Robins, B., Dautenhahn, K., Boekhorst, R.T., Billard, A.: Robotic Assistants in Therapy and Education of Children with Autism: Can a Small Humanoid Robot Help Encourage Social Interaction Skills?. Journal of Universal Access in the Information Society. 4, 105--120 (2005)
21. Smith, C.A., Lazarus, R.S.: Emotion and Adaptation. In: L.A. Pervin (ed.), Handbook of Personality: theory & research, Guilford Press, NY, 609--637 (1990)
22. Ulrich, R.S.: View through a Window May Influence Recovery from Surgery. Science, vol. 224, 420--421 (1984)
23. Ulrich, R.S.: Effects of Interior Design on Wellness: Theory and Recent Scientific Research. Journal of Health Care Interior Design, Vol. 3, 97--109 (1991)
24. Van Vugt, H.C., Hoorn, J.F., Konijn, E.A.: Interactive engagement with embodied agents: An empirically validated framework. Computer Animation and Virtual Worlds, Vol. 20, 195--204 (2009)
25. Verderber, S., Reuman, D.: Windows, Views, and Health Status in Hospital Therapeutic Environments. Journal of Architectural and Planning Research, vol. 4, 120--133 (1987)
26. Wada, K., Shibata, T.: Living with Seal Robots in a Care House - Evaluations of Social and Physiological Influences. In: Intelligent Robots and Systems, IEEE/RSJ International Conference on, pp. 4940--4945 (2006)
27. Wada, K., Shibata, T.: Social Effects of Robot Therapy in a Care House -Change of Social Network of the Residents for One Year-. JACIII. 13, 386--392 (2009)
28. Werry, I. Dautenhahn, K.: Applying Mobile Robot Technology to the Rehabilitation of Autistic Children. In: Proc. of 7th Int. Symp. on Intelligent Robotic Systems, pp. 265--272 (1999)
29. Zwanenburg, E.: Enhancing Believability through Expectation-Based Emotions in a BDI Agent. Master Thesis, Vrije Universiteit Amsterdam, (2009)
30. http://www.few.vu.nl/~ghazanfa/IVA2009/tictactoe.html

PART V EMBODYING EMOTIONS IN VIRTUAL AGENTS

CHAPTER 12

Enhancing Involvement in Financial Services via Intelligent Virtual Agents

Part of this paper will appear as Bosse, T., Siddiqui, G.F., and Treur, J., An Intelligent Virtual Agent to Increase Involvement in Financial Services. In: Proceedings of the 10th International Conference on Intelligent Virtual Agents, IVA'10. Lecture Notes in Computer Science, Springer Verlag, 2010.

Enhancing Involvement in Financial Services via Intelligent Virtual Agents

Tibor Bosse[1], Ghazanfar F. Siddiqui[1,2], and Jan Treur[1]

[1] Vrije Universiteit Amsterdam, Department of Artificial Intelligence,
De Boelelaan 1081a, 1081 HV Amsterdam, The Netherlands
[2] Quaid-i-Azam University Islamabad, Department of Computer Science, 45320, Pakistan
{tbosse, ghazanfa, treur}@few.vu.nl ghazanfar@qau.edu.pk
http://www.few.vu.nl/~{tbosse, ghazanfa, treur}

Abstract. In order to enhance user involvement in financial services, this paper proposes to combine the idea of adaptive personalization with intelligent virtual agents. A computational model for human decision making in financial context is introduced, and is incorporated within an intelligent virtual agent. To test whether the agent enhances user involvement, a web application has been developed, in which users have to make a number of investment decisions. This application has been evaluated in an experiment for a number of participants interacting with the system and afterwards providing their judgement by means of a questionnaire. The preliminary results indicate that the virtual agent can show appropriate emotional expressions related to states like happiness, greed and fear, and has high potential to enhance user involvement.

Keywords: user involvement, finance, greed and risk, adaptive personalization.

1 Introduction

In recent years, there has been a huge increase in the amount of services that are being offered via the Internet. These services include, among others, financial services such as Internet banking [19]. Despite the success of such services, an existing challenge in this area concerns the question how to make people more *involved* in such financial applications. According to [1], customer involvement in financial services can be defined as ‘an unobservable state of motivation, arousal or interest’ (taken from [16]). In order to increase this state of involvement in users of financial applications, some authors claim that personalization is an important criterion (e.g., [2, 7]): by having the system learn certain characteristics of the customer, this person will feel more understood and will be more likely to accept the service that is offered. However, there is also research that suggests that personalization alone is not sufficient for financial services to attract users for longer periods (e.g., [9]).

To deal with this last issue, the current paper proposes to enhance user involvement in financial applications by combining adaptive personalization with the use of an intelligent virtual agent. As pointed out by various authors (e.g., [14]), human-like virtual agents have the ability to increase a user’s presence in virtual environments. This finding was the inspiration to develop a personalized intelligent agent which supports persons that have to make financial (investment) decisions. As known from behavioral economics, humans do not behave completely rationally when they have to decide between alternatives that involve risk (as, for example, in financial situations). Since then, from time to time it has been argued that theories of economic decision making need to incorporate psychological factors such as greed and fear [6, 13, 15, 18]. Thus, the main goal of this paper is to develop a virtual agent that has insight in and adapts to the individual psychological characteristics and states over time of persons that are working with financial applications. The virtual agent should exploit this on the one hand by providing appropriate support, in following these (dynamical) states and characteristics in an adaptive personalized manner. On the other hand, by showing the appropriate emotions at the right moment the virtual agent encourages involvement and reflection by the person through mirroring his or her states and decision making processes; for example, the agent may show the person how greedy he or she behaves.

In order to develop such a supporting virtual agent, as a basis a solid computational model of human decision making in financial context is needed. To this end, the model presented in [3] is taken. This model takes some of the main principles underlying the Modern Portfolio Theory (MPT) [8, 17] as a point of departure, and extends these with mechanisms to incorporate psychological factors (inspired, among others, by [4, 10, 11, 13, 15]). In the current paper, this model is extended and incorporated within an intelligent virtual agent. To test whether the agent enhances user involvement, a simple web application has been developed, in which users make a number of investment decisions. This application has been evaluated by a number of participants in an experiment in which they interacted with the agent and afterwards provided their judgement by means of a questionnaire.

The remainder of this paper is structured as follows. In Section 2, the basic model for financial decision making (taken from [3]) is summarized, and extended to make it suitable for functioning in an interactive context. In Section 3, the extended model is illustrated by means of some simulations. Next, Section 4 describes how the model was incorporated within an intelligent virtual agent. Section 5 introduces the experiment that was performed to evaluate the virtual agent, and Section 6 presents the results. Section 7 concludes the paper with a discussion.

2 The Agent Model to Analyze Financial Decision Making

This section presents the computational model that the virtual agent will use to analyze a human's behavior with respect to financial decision making. In particular, this model will enable the agent to estimate the user's states and characteristics related to greed and risk taking. In Section 2.1, a global overview of the basic model (from [3]) is provided, and in Section 2.2, a mathematical formalization of this model is presented. Section 2.3 presents an extension of the model that allows it to learn the user's personal risk profile.

2.1 Overview

The agent model to analyze the financial decision making of a person is based on the assumption that a person's greed is determined by her (long-term) personality profile (e.g., some persons are more risk seeking than others), combined with observations about recent events (e.g., if many investments have provided high returns recently, persons are more likely to increase their greed, and as a consequence take more risk). These assumptions can also be found in existing literature such as [4, 10, 11, 13, 15]. A global overview of the agent model for financial decision making and its interaction with the world is depicted in Figure 1. In this figure, the dotted boxes represent, from left to right, the human agent that makes the investment decisions, and the world. The small circles denote states of human and world. Note the difference between input states (depicted at the left hand side of a box), internal states (depicted within the box), and output states (depicted at the right hand side). The arrows indicate (causal) relationships between states.

As shown in Figure 1, the agent model includes a notion of greed as a mental state of an individual. This greed is assumed to be a dynamic state, which is continuously influenced by the results of earlier investments and the individual's personality profile concerning risk taking or risk avoidance. The former is considered to be depending, among others on the state of the world's economy, whereas for the moment the latter is assumed to be a static characteristic of an individual. The greed directly affects the investment decisions that the individual makes: a person with higher greed will decide for more risky investments. To create an appropriate economic context, one particular type of investment decision is considered, namely the task to choose a financial product characterized by two factors concerning risk and expected gain. Thus the set of products is represented using a standard risk/return curve as proposed in the literature on Modern Portfolio Theory; e.g., [8].

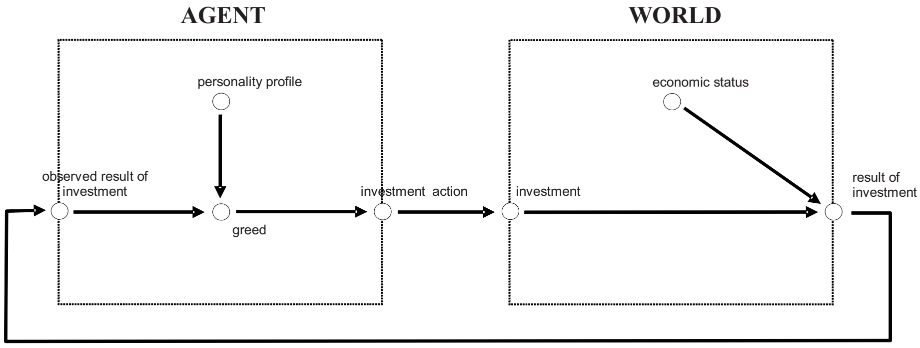


Fig. 1. Overview of the Agent Model and its Interaction with the World

After an investment decision has been made, within the world the results of the investment are determined. These results depend not only on the selected product (in the sense that a more risky product has a lower probability to result in some return), but also on the economic state of the world. For simplicity, the current model considers the economic status as an external variable, although in reality this variable may depend on many other factors, such as the economic behavior of other agents in the system (cf. [4]). The results are observed by the individual, which completes the interactive loop between agent and world.

2.2 Formalization of the Adaptive Decision Making Process Model

The model assumes that the user can choose between a finite number of products (in this paper, 10 products are used). Individuals are able to choose between these products by taking two aspects of a product into account, namely *risk* and *expected return*. To create a realistic range of products, the following parabolic equation is used for the relation between expected risk X and expected return Y (cf. [8, 17]): $X=aY^2+bY+c$ with $a=1$, $b= -0.1$ and $c=0.1$. The graph is shown in Figure 2. The idea is that more greedy persons will select products that are further to the right hand side of the curve.

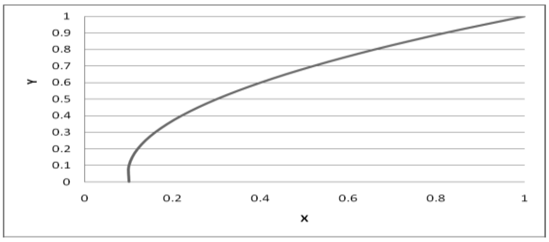


Fig. 2. Expected Risk/Return Curve

In contrast to a risk profile of a human, which is usually considered a static personal characteristic, in the model presented here a dynamic state of greed is used as a basis for decision making. The dynamics of greed are modeled by the following difference equation:

$$G(t+\Delta t) = G(t) + \beta\, [((p+1)/2)\, E(t) - G(t)]\, \Delta t$$

Here $G(t+\Delta t)$ is the updated greed, $G(t)$ is the old greed, $E(t)$ is the world event concerning the return on the earlier chosen product, and p is the individual static risk profile (a personal characteristic; 0 means that the individual is low risk taking and 1 means the individual is high risk taking). The flexibility factor β indicates the proportion of the old greed that is taken into account to determine the new greed; $\beta = 0.1$ has been taken. The values of G, E and p are in the range between 0 and 1. The underlying idea of the formula is that persons may show more greedy behavior if their individual risk profile is more risk taking, and if they have received more positive experiences in the recent past (see also [4, 10, 11, 13, 15]). The initial values of P, G, and E are taken 0.5.

Next, based on the greed and the personality characteristic p the person selects a product. As a first step the following factor r is determined:

$$r = ((1/G)-1\,) \,/\, (2(p+1))$$

This r is taken as the required slope of the curve (depicted in Figure 2) for the product to be chosen, according to Modern Portfolio Theory. The actual choice of the product is made as follows. For each of the considered products (X, Y) the following is calculated: $Z(X, Y) = Y - r*X$. Then the product (X, Y) is chosen with the highest $Z(X, Y)$. This product is the closest approximation of the point at the curve with slope r.

The model for calculating the return E of the investment is as follows; here W is the state of the world (taken between 0 and 1), and (X, Y) is the selected product:

1. *Generate a random number C between 0 and 1 (both inclusive)*
2. *IF* $C \geq X * (1-W)$ *THEN* $E = Y$
3. *IF* $C < X * (1-W)$ *THEN* $E = 0$

This shows, for example, that when the state of the economy W is maximal, there is no risk to have no return, and when W is minimal this risk is with probability X.

The decision model as described is adaptive in the sense that it adapts not only to the state of the world but also to the psychological state of greed of the person. In this way it follows more closely the decision making of a person over time than would be the case when only a static risk profile p was used as a personal characteristic. However, in the dynamics of this state of greed still the risk profile p plays its role. For each user, an appropriate value of this p has to be determined as well. How this is done in an adaptive manner, is described in the following subsection.

2.3 Adapting the Value of the Risk Profile to the User

The risk profile p is a characteristic that depends on the person. In case the model is used only for simulation (as in [3]), the modeler can simply fill in a value for this parameter. However, as in the current paper the model is applied at runtime to real humans, a mechanism is needed to estimate this value in an adaptive manner for a particular individual. To this end, this estimation is based on a person's decisions made. In each iteration, the model compares the actual product as selected by the human (X_h, Y_h) with the product that the software agent predicted to be selected (X_a, Y_a). For each of these products (X, Y) the value of $r(X, Y)$ is determined as indicated in the previous section. Then the value for p is adapted as follows.

$$p(t+\Delta t) = p(t) + \eta\, p(t)\, [\, (r(X_a, Y_a) - r(X_h, Y_h))/\, r(X_h, Y_h)\,]\, \Delta t \qquad \text{if} \quad r(X_a, Y_a) - r(X_h, Y_h) \leq 0$$

$$p(t+\Delta t) = p(t) + \eta\, (1-p(t)\,)[\, (r(X_a, Y_a) - r(X_h, Y_h))/\, r(X_h, Y_h)\,]\, \Delta t \qquad \text{if} \quad r(X_a, Y_a) - r(X_h, Y_h) \geq 0$$

Here the adaptation rate η has been chosen 0.8. Thus, this mechanism adapts the value for p with a percentage that is proportional to the difference between the actual and predicted value of r.

3 Example Simulations

A number of simulation experiments have been performed before incorporating the model in the virtual agent application. First, some experiments were performed to evaluate the global behavior of the model. An example simulation trace is shown in Figure 3. Here, time is on the horizontal axis and G, W, E and the average profit received by the person are represented on the vertical axis. Moreover, the initial value of $G = 0.5$, $p = 0.6$, and $\beta = 0.1$. For the value of W, a scenario has been established that is based on existing empirical data. For these data, the global Gross Domestic Product data over the period 1969-2008 have been taken from [21]. As illustrated by this simulation, the person continuously adapts its greed G to the status of the economy W. Thus, the person quickly learns which level of greed is most appropriate in which situation. For more simulation results, see [3].

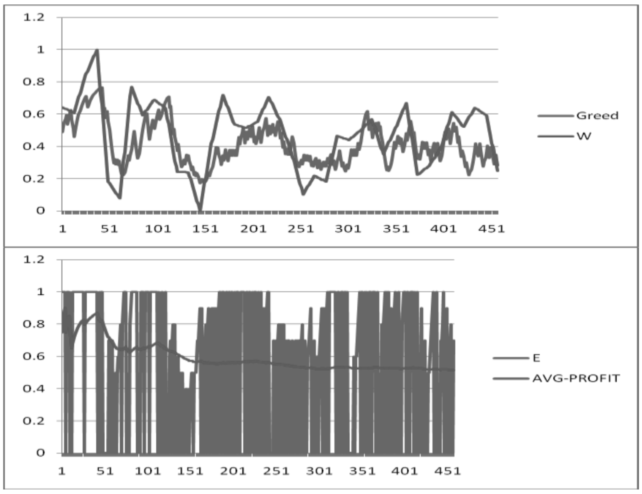
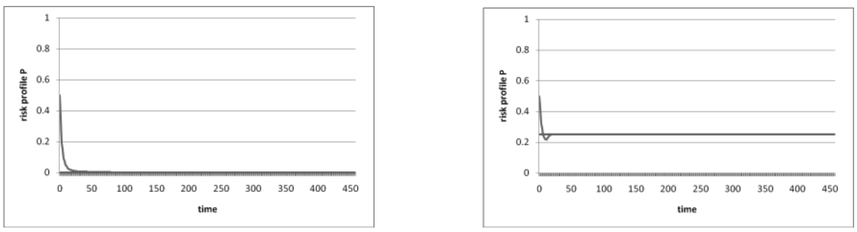


Fig. 3. Example Simulation Results

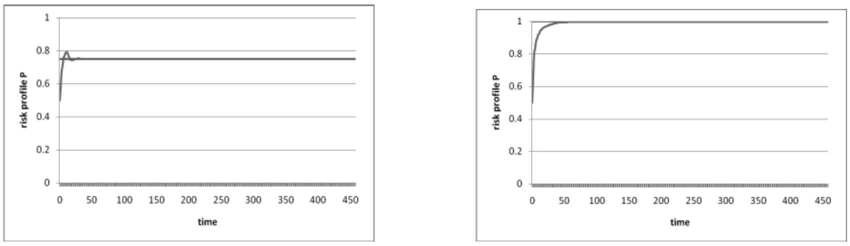
Next, some simulation experiments have been performed to test whether the agent correctly learns the risk profile of a (simulated) person. To this end, two versions of the decision making model were simulated: one for the human, and one for the agent which has to estimate the human's risk profile. Both models used the same formulae and parameter settings, but the values for p were different. For the human, different (static) values for p were taken, as shown in Figure 4: $p_h = 0$, $p_h = 0.25$, $p_h = 0.75$ and $p_h = 1$. For the agent, as initial value, $p_a = 0.5$ was taken, but over time this value was adapted based on the mechanism described in Section 2.3. Moreover, in these formulae, the adaptation rate η was taken 0.8.

As indicated by the different graphs in Figure 4, in all cases the model was able to find the person's risk profile relatively quickly (at most within about 50 iterations). In some cases the model over-compensates a bit due to the high value of η, but this effect is minimal.

Despite this positive result, one should be aware that these were ideal scenarios for the model to learn the human's risk profile, since the (simulated) humans behaved exactly according to the decision making model. For real humans, this behavior will be more irregular, and thus the risk profiles will be more difficult to learn.



$p_h = 0$ $\qquad\qquad$ $p_h = 0.25$



$p_h = 0.75$ $p_h = 1$

Fig. 4. Risk Profile of Agent and Person (Agent: blue line, Person: red line)

4 The Use of a Virtual Agent

The agent model described in Section 2 is able to analyze the human's decision making by observing her decisions and the received returns, while tuning the risk profile to the person. Within this analysis not only this personal risk profile is available, which is assumed static for the person, but also the more dynamic greed level that actually determines the decisions. By having this, at each point in time the model can predict what a reasonable decision would be for the human, given her personal background and history. In particular at all stages of the process it can estimate and show the type and level of emotions expected. These emotions have been incorporated within a virtual agent, which can be shown to the human at runtime.

To design and implement the virtual agent, Haptek's Peopleputty software [22] has been used. Through this software the face of the virtual agent was created. More specifically, twelve different faces were designed using the built in sliders for happy, sad, anger, mellowness, suspicion, and curiosity (which are related to facial expressions), and ego, aggression and energy (which are related to head movement). Each of these twelve faces represented a particular emotional state; one for each possible combination of the three required levels of the emotions *happiness* (slightly_happy, happy and very_happy), *sadness* (slightly_sad, sad and very_sad), *fear* (slightly_scared, scared and very_scared) and *greed* (slightly_greedy, greedy and very_greedy). In addition, a face for the state neutral was used. A web-based application was implemented, within which the virtual agent was embedded as a Haptek player. For this the scripting language JavaScript [20] was used, in combination with scripting commands provided by the Haptek software [22], to control the Haptek player within a web browser.

Within the application, a human can make a number of consecutive investment decisions, while the virtual agent mimics the estimated emotional states related to happiness, sadness, greed and fear of the human (see the screenshot in Figure 5). In this application, in total 10 (represented by letters from *A* to *J*) products are given. The characteristics of these products are represented by the two variables *X* and *Y*, which are shown on the right hand part of the screen. Here, as in Section 2, *X* represents the risk of the product, and *Y* represents the expected return of the product. Note that in the model both *X* and *Y* have a value in the domain [*0, 1*], but in the application the values of *Y* have been scaled to the domain [*0, 1000*], to have them represent US dollars.

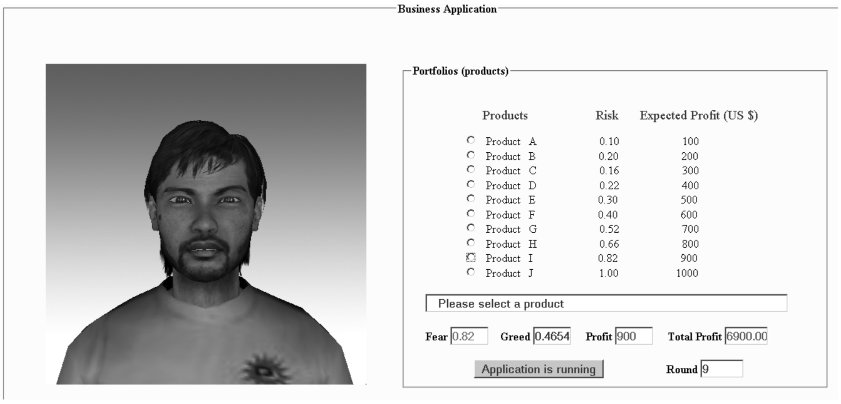


Fig. 5. Screenshot of the Application

During a number of rounds, the human is asked to select a product from the given products (from A to J). After she selected a product, some time will pass, until a message is shown on the screen that the "result of your investment will soon be announced". Next, it again takes some seconds until the real result is shown on the screen. To determine what this result will be, the formulae shown in Section 2.2 are used. Note that these formulae make use of the parameter W which represents the economic situation of the world. The value of this parameter fluctuates over time, and is not known to the user.

In every round, the virtual agent shows emotional facial expressions at appropriate moments. The following fixed scheme determines when to show which type of emotional expression:

1. human is asked to select a product
2. agent shows face related to greed
3. human selects a product
4. agent shows face related to greed
5. message is shown stating that result will soon be announced
6. agent shows face related to fear
7. result of investment is announced
8. agent shows face related to happiness/sadness
9. go back to 1.

The criteria that determine the exact faces that are displayed are as shown in Table 1 (where risk equals the X value of the selected product, profit equals the result of the investment (i.e., either 0 or the Y value of the selected product), and greed equals the value of G as estimated by the model).

When the agent shifts from one facial expression to another, it would be undesirable if the emotions of the agent would shift too quickly. Therefore, these shifts are performed in a more fluent manner. For instance, if the agent shifts from very happy to very greedy, the following faces are shown consecutively:

very happy → happy → slightly happy → neutral → slightly greedy → greedy → very greedy

Such a scenario is used when the agent shifts from any emotional state to another emotional state.

Table 1. Criteria for the displayed Face Expressions

	Criterion	*Displayed*
Fear	risk ≤ 0.5	neutral face
	risk > 0.5 & risk ≤ 0.7	scared face
	risk > 0.7 & risk ≤ 1	very scared face
Happiness/Sadness	profit = 0	very sad face
	profit > 0 & profit ≤ 300	slightly happy face
	profit > 300 & profit ≤ 600	happy face
	profit > 600) & profit ≤ 1000	very happy face
Greed	greed ≤ 0.1	neutral face
	greed > 0.1 & greed ≤ 0.3	slightly greedy face
	greed > 0.3 & greed ≤ 1	very greedy face

While the application is running, some information about the user is displayed in the bottom right part of the screen (see Figure 5). This information concerns the user's estimated amount of fear and greed (in the domain [*0, 1*]), her current amount of profit received, and her total (cumulative) amount of profit.

5 Experiments with the Virtual Agent

A number of experiments were performed to test to what extent users of the application feel involved with the agent. In total, 15 participants were recruited to perform the experiment. The age of the participants ranged between 24 and 34, with a mean age of 29 and a standard deviation of 2.78. Among the participants, 11 were male and 4 were female.

Two variants of the experiment were designed, one with which the virtual agent was showing the appropriate emotions and one in which it did not show any emotions. All of the participants were asked to perform both variants (where we used counterbalancing to determine the order in which they were performed).

Before they started the experiment, each participant was first are asked to read the following instructions:

> *Imagine that you are an investor in a stock market. During a number of subsequent rounds, you have to select a product from a given set of products. Each round, the same 10 products are available. The characteristics of these products are represented by two variables (called X and Y), which are shown on the screen. Here, X is a value in the domain [0, 1] which represents the* ***risk*** *of the product (i.e., a higher value for X means that it is more likely that you will* ***not*** *receive the corresponding return), and Y is a value in the domain [0, 1000] which represents the* ***expected return*** *of the product in US dollars (i.e., a higher value for y means that you will earn more profit). The value of X is related to the probability p of* ***not*** *receiving the corresponding return Y according to the formula p = X * (1-W). Here, W is a number in the domain [0, 1] which represents the economic situation of the world (i.e., a higher value for W means that it is more likely that you will receive the corresponding return). However, the value of W fluctuates during the simulation, and is not shown to you. After you have selected a product, some time will pass, until a message is shown on the screen that the result of your investment will soon be announced. Next, it still takes some seconds until the real result is shown on the screen. As mentioned above, the probability of receiving the profit also depends on the economic status of the world. After the result of your investment has been announced, a new round starts, in which you are asked to make a new investment. In total, the experiment lasts 20 rounds.*

Next, a small training was given to each participant, and after that the participants performed the actual experiment. When the experiment was finished, the person was asked to fill in a questionnaire. In this questionnaire (cf. [5]), the participants were asked to evaluate, using a 7 point Likert scale [12] (with 1=strongly disagree, 7=strongly agree and 4=neutral), various properties of the agent related to involvement. In particular, they were asked whether they thought the virtual agent was friendly, trustworthy, showing emotions adequately, realistic, showing happiness, showing sadness, showing greed, showing fear and human-like.

In each experiment, the value of the economic state *W* fluctuated between *0* and *1*, as shown in Figure 6. However, the participants were not aware of this.

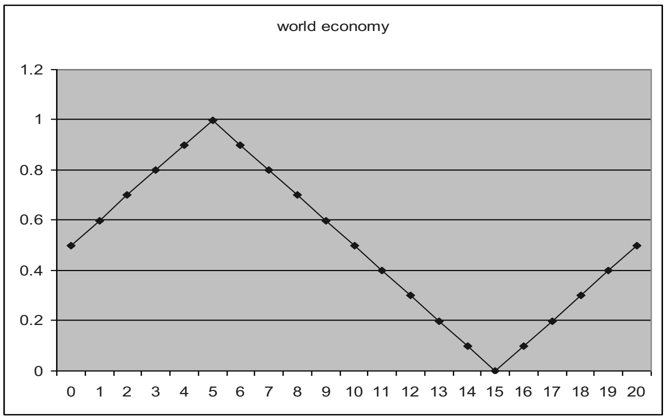


Fig. 6. Fluctuation of *W* during the Experiment

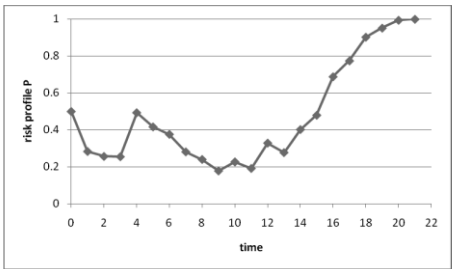
6 Results of the Experiment

This section presents the results of the experiment. First, some results are shown with respect to the agent's ability to learn the user's risk profile. After that, the results of the questionnaire are discussed.

6.1 Risk Profile

During the experiments, the values of the selected products, the person's estimated risk profile *p*, and greed value *G* have been logged. This allows us to investigate the dynamics of the person's characteristics as estimated by the agent. In order to illustrate this, we consider two scenarios.

The risk profile *p* of two different persons as learned by the agent during the experiment is shown in Figure 7 and 8. In these figures, time is along the x-axis whereas the value of the estimated risk profile *p* is along the y-axis. In Figure 7, the agent interacts with a person with an extremely high risk profile, as this person was always selecting products with a high risk value. As can be seen from Figure 7, initially the value of estimated value of *p* fluctuates a lot, but eventually this value approaches 1. In this scenario, the person eventually obtained a profit of 11800 US$.



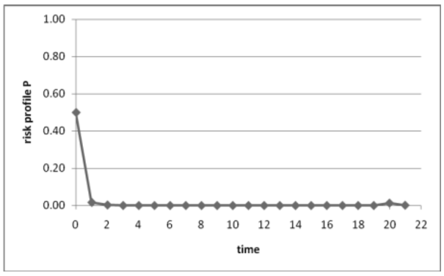


Fig. 7. Risk Profile of Person 1

Fig. 8. Risk Profile of Person 2

In Figure 8, the agent interacts with a person with an extremely low risk profile. In this scenario, the person was always selecting products with low risk. As can be seen from Figure 8, initially the value of risk profile p drops to 0.017, then to almost 0 and stays around this value for the rest of the scenario. In this scenario, the person eventually obtained a profit of 2900 US$.

In addition, Figure 9 shows the risk profiles of 6 different persons (P1 to P6) in one picture. As can be seen from this figure, P6 is very conservative, P2 is moderately conservative, P1 and P3 are moderately aggressive, and P4 and P5 are very aggressive investors.

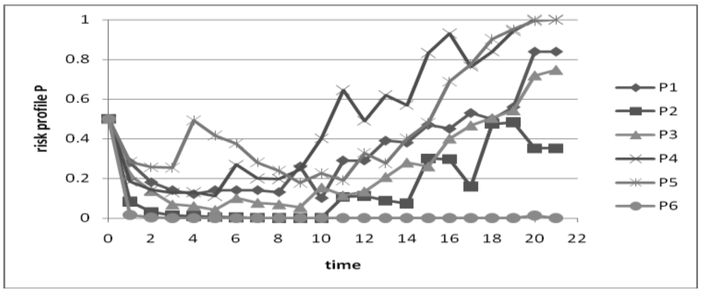


Fig. 9. Risk Profile of 6 Different Persons

In the above experiments, the users did not know about the state of the world economy while making their decisions. To test whether such knowledge would make a difference, some additional experiments (with 6 new participants) have been conducted, in which the users are made more aware of the economic situation during the scenario (i.e., at each round the value of W, in the domain $[0, 1]$, is shown to the user). Two different scenarios for the experiments have been created, one with a fluctuating value for W (identical to the scenario used above, see Figure 6), and one in which W was constantly set to 0.5. The results of these experiments are summarized in Figure 10 and 11. In the latter case, with a constant world state the values seem to stabilize more than in the other cases. These results may suggest that the person's level of awareness of the world state may be a nontrivial factor in this context.

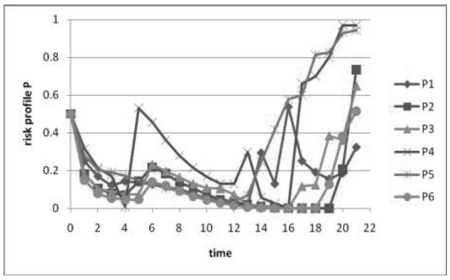


Fig. 10. Risk Profile during Scenario 1 (fluctuating world economy)

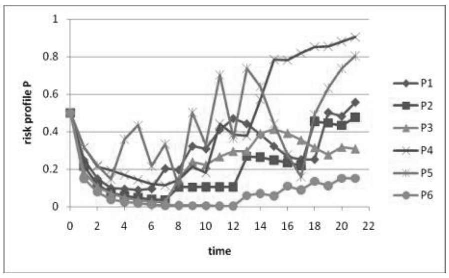


Fig. 11. Risk Profile during Scenario 2 (constant world economy with *W*=0.5)

6.2 Questionnaire

In addition, the answers provided by the participants to the questions about their involvement with the virtual agent were analyzed by means of paired sample t-tests. The results are shown in Table 2.

Table 2. Results of the Questionnaire

Q #	Question	With Emotions		Without Emotions		Paired Sample Test	
		Mean	SD	Mean	SD	t	Sig(2tailed)
1	Friendly	4.47	1.506	3.53	1.992	2.168	.048
2	Trustworthy	4.13	1.598	2.47	1.246	4.183	.001
3	Adequate emotions	4.93	1.223	2.40	1.844	4.219	.001
4	Realistic	4.93	1.534	3.47	1.922	2.442	.028
5	Happiness	5.84	0.834	2.13	1.457	9.153	.000
6	Sadness	5.67	1.113	2.20	1.474	7.124	.000
7	Greed	2.67	1.877	2.00	1.195	1.323	.207
8	Fear	4.27	1.751	2.00	1.254	5.264	.000
9	Happiness at right time	5.60	0.986	2.20	1.656	7.462	.000
10	Sadness at right time	5.87	0.990	2.00	1.363	9.648	.000
11	Fear at right time	4.13	1.598	1.93	1.387	4.036	.001
12	Human-like	4.67	1.291	2.93	1.667	2.578	.022

As shown in the table, for all properties except 'greed', the participants scored the virtual agent with emotions above moderate. Similarly, for all properties except 'greed', the participants appreciated this variant more than the virtual agent without emotions. The participants were also asked to give suggestions or comments about the application. Some participants said, for example, that the fear emotion should be more intense, while others said that the greed emotion should be improved, as they did not see this very well. Participants also indicated that they were more involved with the virtual agent with emotions. In addition, some participants were of the opinion that the agent should speak as well.

The fact that greed did not score very well in this first test may depend on the type of face expression chosen for greed. To explore how the perception of greed could be enhanced, another small experiment was conducted. For this purpose 9 different alternative faces for greed were created using the Peopleputty software [22]. Six new participants were asked to rate each face on a 7 point scale, for its appropriateness to express greed. After all the participants gave their responses, for each face the average score over all participant responses was taken, and the face with the highest average value was selected for a next experiment (again with 6 new participants). This experiment was identical to the experiment of which the results are given in Table 2, only in this case the new face was used to display the greed. Table 3 shows (part of) the results of this experiment.

Table 3. Results of the Questionnaire for the Additional Experiment related to Greed

Q #	Question	With Emotions		Without Emotions		Paired Sample Test	
		Mean	SD	Mean	SD	t	Sig(2-tailed)
1	Greed	5.167	0.983	2.33	1.366	3.782	.013

As can be seen, this time the greed was evaluated much better than in the first experiment, and resembles the evaluations of the other emotions.

7 Discussion

In this paper an adaptive agent model was presented combining adaptive personalization with intelligent virtual agents, in order to enhance user involvement in financial services. To this end a computational model for human decision making in financial context was introduced, and incorporated within an intelligent virtual agent. This computational model enables the virtual agent to have a form of understanding of the person's (dynamical) states and decision making processes in an adaptive manner. Moreover, a second way in which the agent was made adaptive was by equipping it with a model to tune the risk profile parameter to the person.

A web application has been developed, in which users make a number of investment decisions. This application has been used to test whether the virtual agent enhances user involvement. This has been evaluated in an experiment for a number of participants interacting with the system and afterwards providing their judgement by means of a questionnaire. The preliminary results indicate that the virtual agent can be given facial expressions showing emotional states like happiness, greed and fear that are evaluated as appropriate (regarding the type of facial expression as well as the moments on which the expressions are shown). In particular, getting an appropriate expression for greed was nontrivial.

For future work the virtual agent may be tested in a real environment to analyze whether it makes humans perform better in financial decision making, for example in the form of a Smartphone application. One of the factors that may need some more attention is the level of awareness of the person of the state of the world.

References

1. Aldlaigan, A.H. and Buttle, F.A. (2001). Consumer involvement in financial services: An empirical test of two measures. *International Journal of Bank Marketing, 19*(6), pp. 232-246.
2. Blom, J. and Monk, A. (2001). One-to-one e-commerce: who's the one? In: *Proceedings of the Conference on Human Factors in Computing Systems (CHI'01)*. ACM Press, pp. 341-342.
3. Bosse, T., Siddiqui, G.F., and Treur, J. (2010). A Personalized Intelligent Agent Model for Financial Decision Making Incorporating Psychological States for Greed and Risk. Technical Report, Vrije Universiteit Amsterdam.
4. Bosse, T., Siddiqui, G.F., and Treur, J. (2010). Modeling Greed of Agents in Economical Context. In: *Proceedings of the 23rd International Conference on Industrial, Engineering & Other Applications of Applied Intelligent Systems, IEA/AIE'10*. Lecture Notes in Artificial Intelligence, Springer Verlag, 2010, in press.
5. Bowling A. (1997). Questionnaire design. In: *Research Methods in Health*. Buckingham: Open University Press, 1997.
6. Brandstätter, H. (1993). Should Economic Psychology Care about Personality Structure? *Journal of Economic Psychology*, vol.14, pp. 473-494.
7. Chau, P.Y.K. and Lai, V.S.K. (2003). An empirical investigation of the determinants of user acceptance of internet banking. *Journal of Organizational Computing & Electronic Commerce*, 13(2), pp. 123-145.
8. Haugen, R.A. (1997). *Modern Investment Theory*. Prentice-Hall, Inc.
9. Ho, S.Y. (2006). The Attraction of Internet Personalization to Web Users. *Electronic Markets*, vol. 161, pp. 41-50.
10. Kasser, T., Cohn, S., Kanner, A., Ryan, R. (2007). Some costs of American corporate capitalism: A psychological exploration of value and goal conflicts. *Psychological Inquiry*, vol.18, pp.1-22.
11. Krawiec, K.D. (2000). Accounting for Greed: Unravelling the Rogue Trader Mystery. *Oregon Law Review*, vol 79, issue 2, pp. 301-339.
12. Likert, R. (1932). A Technique for the Measurement of Attitudes. *Archives of Psychology*, vol. 140, pp. 1-55.
13. Lo, A., Repin, D.V., and Steenbarger, B.N. (2005). Fear and Greed in Financial Markets: A Clinical Study of Day-Traders. *American Economic Review* 95, pp. 352-359.
14. Nowak, K.L. and Biocca, F. (2003). The effect of the agency and anthropomorphism on users' sense of telepresence, copresence, and social presence in virtual environments. *Presence: Teleoperators and Virtual Environments 12*(5), pp. 481-494.
15. Rabin, M. (2002). A Perspective on Psychology and Economics. *European Economic Review*, vol. 46, pp. 657-685.
16. Rothschild, M.L. (1984). Perspective on Involvement: Current Problems and Future Directions. In: Kinnear, T.C. (ed.), *Advances in Consumer Research*, Association for Consumer Research, vol. 11.
17. Sabal, J. (2002). *Financial Decisions in Emerging Markets*. Oxford University Press, Inc., New York.
18. Simon, H.A. (1987). Behavioral Economics. In: *The New Palgrave: A Dictionary of Economics*. London, MacMillan.
19. Tan, M. and Teo, T.S.H. (2000). Factors Influencing the Adoption of Internet Banking. *Journal of the AIS* 1(5), pp. 1-42.
20. http://developer.mozilla.org/en/docs/About_JavaScript
21. http://www.ers.usda.gov/Data/Macroeconomics/
22. http://www.haptek.com

PART VI DISCUSSION AND FUTURE WORK

CHAPTER 13

Discussion and Future Work

Discussion and Future Work

The main goal of this dissertation as described in the first chapter is to explore how computational models of affect can be integrated within virtual agents. To this end different models have been analyzed, formalized, combined, simulated, and evaluated within applications in health care, business and game context. This thesis is composed of six parts, which comprise in total 13 chapters. For modeling and implementation, different modeling environments and implementation techniques were exploited, e.g., LEADSTO, C++, HTML, JAVASCRIPT, PHP and HAPTEK PeoplePutty. To check whether the models behaved as expected, both simulation experiments and empirical validation (partially) have been performed.

Overview

The research presented in each part of the thesis is discussed below.

In **Part II**, the first investigated model put forward describes to which extent an agent becomes involved with another agent, or stays at a distance from it, by formalizing in a computational manner the previously informally described I-PEFiC model. More specifically, in **Chapter 2**, the informal theoretical model for involvement-distance trade-offs presented in [14] has been translated into a computational model, formalized in the language LEADSTO [1]. Two main results were established. First, the model turned out to be adequate for simulating the dynamics of involvement distance trade-offs and their influence on satisfaction. Second, and perhaps more surprisingly, it was found that positive features can increase the level of distance, and that negative features can increase involvement. This is explained by the fact that these aspects do not directly influence involvement and distance, but only indirectly via the factors of similarity, relevance and valence. Although this finding may be counterintuitive, it corresponds to empirical evidence presented in [25, 26]. **Chapter 3** presents an extension of the computational I-PEFiC model as described in Chapter 2 by goal-directed judgment formation and overt actions to enable software agents to combine rational with affective processing. Models of decision-making usually assume the process to be rational, which would exclude the possibility of emotions playing a role other than by disturbing the process [12]. However, humans often involve the emotions they feel in decision making in a constructive manner [24]. Simulation experiments have been conducted, and it was found that agents preferred affect-driven decision options to rationally driven decision options in situations where choices for low expected utility are irrational. **Chapter 4** is an extension of the work described in Chapter 3, which manages that the actions agents undertake have an effect on other agents. The agents change their perceptions and beliefs about other agents if actions are taken.

In **Part III**, the integration of three models of affect is addressed. The approach taken in this part was to select three of the more influential models, which share that they can be used to enhance believability of virtual characters: CoMERG [2] (Computational Model of Emotion Regulation based on Gross theory), EMA [20] (Emotion and Adaptation) and I-PEFiC^{ADM} [16] (i.e., the model developed in Part II). First, in **Chapter 5** these three computational models, which describe the processes related to emotion elicitation and regulation, are compared. The theories by which they were inspired cover a large amount of psychological literature in affect-related processes, including the works of Frijda [8], Lazarus [18], and Gross [10]. In this chapter, it is argued that each of the three approaches has its specific focus. For example, CoMERG [2] covers a wide variety of emotion regulation strategies, whereas I-PEFiC^{ADM} [16] provides an elaborated mechanism for encoding of different appraisal domains, which have empirically shown to be crucial in human-robot interaction. EMA on its turn contains very sophisticated mechanisms for both appraisal and coping, which have already proved their value in various applications. Because several of these features are complementary to each other, this chapter explores possibilities to integrate them into one combined model of affect for virtual agents. **Chapter 6** presents Silicon Coppélia, a computational model which integrates of CoMERG [2], I-PEFiC^{ADM} [16], and EMA [20]. Compared to the model in I-PEFiC^{ADM} [16] the agents have goal-related beliefs that lead to emotions. Also, some emotion regulation strategies were added to the system based on [2, 24]. Simulation experiments show that Silicon Coppélia can simulate richer agent behavior than CoMERG [2], I-PEFiC^{ADM} [16] or EMA [20] can do alone.

Part IV focuses on affective states of a person in financial/economical context. This part addresses how involvement in social life in the form of individual investment decisions depends on a personal risk profile, the state of greed of the person, and the state of the world economy. **Chapter 7** discusses similarities and dissimilarities between agent-based models and population-based models in this context. Inspired by variants of predator-prey models (e.g., [5, 19, 21, 27]), a dynamic model was developed for the relationship between individual agents' greed and the state of the global economy. Simulation experiments for different population sizes were performed for both an agent-based and a population-based model. It turned out that, in particular for not too small population sizes, the differences in the economy and average greed between agent-based and population based simulations are close to zero. **Chapter 8** presents a different agent-based model of human financial decision-making behavior in economic context, based on psychological states and characteristics concerning greed and risk taking or risk avoidance. The model takes into account ideas underlying the Modern Portfolio Theory [13, 22] within finance, and incorporates a psychological concept greed and a risk characteristic. To evaluate the model a number of simulation experiments have been performed, which illustrate the model's ability to mimic investment behavior depending on the types of personality and the state of the economy.

Part V addresses applications of virtual agents that show emotions. Here some of the computational models as presented in earlier parts are built in virtual agents in order to let them show the right emotions at the right time. **Chapter 9** aims at building in emotion regulation into virtual characters. To this end, the informal model by Gross [10] was taken as a basis, as previously formalized using a dynamical system-style modeling approach [2]. A virtual environment has been created, which includes a number of virtual agents that have been equipped with the formalized model for emotion regulation. To test the behavior of the model in a prototyping phase, a series of simulation experiments has been performed using the LEADSTO simulation environment [1]; using the Vizard Virtual Reality Toolkit [28], these simulations have been visualized in a graphical environment. The resulting movies provide a first indication that the emotion regulation strategies as described by [10] have been implemented successfully within the virtual characters. The simulation results have been compared with the behaviors for different situations as described by Gross [10] and found consistent. **Chapter 10** presents a virtual agent that guides a person through the Beck Depression Inventory (BDI) questionnaire, which is used to estimate the severity of a depression. The agent responds empathically to the answers given by the user, by changing its facial expression. This resembles face to face therapy more than existing web-based self-help therapies. **Chapter 11** describes a study of an affective virtual agent that can play tic-tac-toe. Being equipped with the integrated model as discussed in Chapter 6, which is an integration of three affect-related models as discussed in Chapter 5, it can show human-like emotional behavior. Simulation experiments with a number of agents were performed under different parameter settings, and pointed out that the agent is able to indeed show human-like emotional behavior, and can make decisions based on rational as well as affective influences. **Chapter 12** deals with a web-based financial investment support system equipped with a virtual agent. The virtual agent in the application tries to replicate human emotions, for example, related to greed, fear and disappointment when a person makes investment decisions and learns about the returns. Experiments with humans indicated that the virtual agent enhanced the user's involvement in the application.

Characterization of the chapters

Various similarities and differences in characteristics can be found when comparing the different chapters of the thesis. In Table 1, an overview of the characteristics of the different chapters of the thesis is presented. The table indicates which chapter uses which theories or models, the domain of the model, the modeling environment for implementation used, and whether simulation and/or empirical validation was performed. Column 1 and 2 represent part and chapter numbers respectively, the next column indicates which theories or models were used as inspiration for the corresponding chapters. Column 4 indicates the discipline of the domain of the model (psychology, social science, economics). Next to this, Column 5 indicates which modeling environments were used in which chapters. Column 6 and 7 indicate whether simulations and/or empirical validation were used in each of the different chapters. A cross indicates that a certain dimension is employed in a certain chapter. The abbreviations used in Table 1 are given below.

P	= Psychology	S	= Simulation
EF	= Economics and Finance	SS	= Social Sciences
EV	= Empirical Validation	MPT	= Modern Portfolio Theory
FTE	= Frijda, Theory of Emotion	PBM	= Population Based Modeling
EMA	= Emotion and adaptation		
GTER	= Gross Theory of Emotion Regulation		
LBE	= Literature from Behavioral Economics		
EKCP	= Expert Knowledge from Clinical Psychologists		

I-PEFiC = Interactively, perceiving and experiencing fictional characters

CoMERG = Computational Model of Emotion Regulation based on Gross theory

Table 1. Overview of the aspects addressed in different chapters of the thesis

Part No.	Chapter No.	Theories OR Models used	Model Domain	Modeling Environment	S	EV
II	2	I-PEFiC	SS P	LEADSTO	X	
	3	I-PEFiC FTE	SS P	LEADSTO	X	
	4	I-PEFiC, I-PEFiC^{ADM} FTE	SS P	LEADSTO	X	
III	5	CoMERG EMA I-PEFiC^{ADM}	SS P			
	6	CoMERG EMA I-PEFiC^{ADM}	SS P	JavaScript	X	
IV	7	LBE	EF P	C++	X	
	8	MPT LBE	EF P	LEADSTO C++	X	
V	9	GTER CoMERG	P	Vizard VR Toolkit	X	
	10	EKCP	P	JavaScript	X	X
	11	CoMERG EMA I-PEFiC^{ADM}	SS P	JavaScript	X	
	12	MPT LBE P	EF P	JavaScript	X	X

As can be seen from Table 1, most models and theories used as inspiration were taken from Social Sciences and Psychology, except Chapters 7, 8 and 12 in which literature from Economics and Finance was taken as a point of departure. In this thesis, an interdisciplinary approach was taken; to this end different theories / models were taken from different fields and have been combined to develop integrative models of affect. Most of these models have first been used for simulation, to test whether their overall behavior was satisfactory. For this, various modeling environments have been used, such as LEADSTO and C++. Next, in some chapters these models have been incorporated within applications related to health care, game and business context. These applications have been developed using JavaScript, in combination with the Vizard toolkit. Finally, for two of the applications (Chapter 10 and 12), user tests have been performed. The preliminary results of these tests show that through these applications, users are more involved in the applications.

Future Work

Concerning future work, the work discussed in the thesis opened numerous possibilities. In this section a number of them are discussed.

A first direction for future work would be to validate the model discussed in Part III. For this purpose a speed date application will be developed in which users will interact with the agent to make an appointment for a date. The dating partner would be performed by an I-PEFiC^{ADM} [16] based software agent. The emotionally laden context of speed dating will be chosen because that context will make it easy to ask the users what the invisible counterpart would think of them, ethically, aesthetically, and whether they believe the other would want to make an appointment with them. Attention will be paid to five issues of I-PEFiC^{ADM} [16] that are particularly of interest to a speed date situation. The issues are Aesthetics (beautiful, ugly), Ethics (good, bad), relevance to personal concerns [9] (Relevance), feeling involved (Involvement), feeling distance (Distance), and intentions to use e.g., willingness to meet again (Use Intentions). An interesting question is whether users would recognize that the agent is making ethical and aesthetic assessments and that these assessments were affecting the agent's level of involvement with them as a dating partner as well as the agent's intentions to use (i.e. meet) them again either in another dating session or in real life. During these speed-dating sessions possibly 7 topics (Family, Sports, Appearance, Hobbies, Music, Food, and Relationships) will be used. Using the application, various hypotheses as derived from the I-PEFiC^{ADM} [16] model will be tested.

Another direction for future work would be to extend the application discussed in Chapter 12. The idea is to develop a banking application in which a virtual agent guides a client during investment decisions. The application would be equipped with an intelligent virtual agent (IVA), which will determine the risk profile of the client, among others using a questionnaire. After that, the IVA will decide, what kind of investment is preferable for the client. In addition to that the IVA will also try to regulate the emotions of the client. The application would be developed in JavaScript and PHP and the IVA would speak with the agent interactively.

A final possibility for future research is to extend the application discussed in Chapter 10. As a first step, it is planned to validate the application with patients with depression. After validation, similar kinds of self-help therapy applications for Cognitive Behavioral Therapy (CBT) and Reinforcing Behavioral Therapy (RBT) can be developed, as it is evident from literature that CBT is effective for the treatment of a variety of problems, including mood, anxiety, personality, eating, substance abuse, and psychotic disorders [6], [7]. By developing such applications, the potential of the developed models for affect would be further explored.

References

1. Bosse, T., Jonker, C.M., Meij, L. van der, and Treur, J. (2007). A Language and Environment for Analysis of Dynamics by Simulation. International Journal of Artificial Intelligence Tools. Vol. 16, 3, 2007, pp. 435-464.
2. Bosse, T., Pontier, M., and Treur, J., A Computational Model based on Gross' Emotion Regulation Theory. Cognitive Systems Research Journal, 2010, to appear
3. Bosse, T., Zwanenburg, E.: There's Always Hope: Enhancing Agent Believability through Expectation-Based Emotions. In: Pantic, M., Nijholt, A., Cohn, J. (eds.), Proceedings of the 2009 International Conference on Affective Computing and Intelligent Interaction, ACII'09, 111-118, IEEE Computer Society Press (2009).
4. Breazeal, C.: Emotion and sociable humanoid robots. In E. Hudlika (Ed.), International Journal of Human Computer Interaction, Vol. 59, pp. 119-155 (2003).
5. Burghes, D.N., Borrie, M.S.: Modeling with Differential Equations. John Wiley (1981)
6. Butler AC, Chapman JE, Forman EM, Beck AT (January 2006). "The empirical status of cognitive-behavioral therapy: a review of meta-analyses". Clin Psychol Rev 26 (1): 17-31.
7. British Association for Behavioral and Cognitive Psychotherapies: What are Cognitive and/or Behavioral Psychotherapies? Retrieved on 2008-11-1
8. Frijda, N. H.: The Emotions. New York: Cambridge University (1986)
9. Frijda, N.H., & Swagerman, J. (1987). Can Computers Feel? Theory and Design of an Emotional System. Cognition and Emotion, 1, 235-258.
10. Gross, J.J.: Emotion Regulation in Adulthood: Timing is Everything. Current directions in psychological science, Vol. 10(6), pp. 214-219 (2001)
11. Gmytrasiewycz, P.J., and Lisetti, C.L. "Emotions and Personality in Agent Design and Modeling" In: M. Bauer,
P.J. Gmytrasiewicz, and J. Vaassileva (Eds.), Springer-Verlag Berlin Heidelberg , 2001, PP. 237-239.
12. Gutnik, L.A., Hakimzada, A.F., Yoskowitz, N.A., and Patel, V.L. "The role of emotion in decision-making: A cognitive neuroeconomic approach towards understanding sexual risk behavior" In: Journal of Biomedical Informetics, volume 39, 2006.
13. Haugen, R.A. (1997). Modern Investment Theory. Prentice-Hall, Inc.
14. Hoorn, J. F. (2008). A Robot's Experience of its User: Theory. In: Sloutsky, V., Love, B.C., and McRae, K. (eds.), Proceedings of the 30th International Annual Conference of the Cognitive Science Society, CogSci'08, 2008
15. Hoorn, J.F., Konijn, E.A., Van Vliet, H., and Van der Veer, G. "Requirements change: Fears dictate the must haves; desires the won't haves" In: Journal of Systems and Software, volume 80, 2007, pp. 328-355.
16. Hoorn, J. F., Pontier, M., and Siddiqui, G. F., When the User Is Instrumental to Robot Goals: First Try – Agent Uses Agent. In: Jain, L., Gini, M., Faltings, B.B., Terano, T., Zhang, C., Cercone, N., Cao, L. (eds.), Proceedings of the 8th IEEE/WIC/ACM International Conference on Intelligent Agent Technology, IAT'08. IEEE Computer Society Press, 2008, pp. 296-301.
17. Konijn, E.A., and Hoorn, J.F.: Some like it bad. Testing a model for perceiving and experiencing fictional characters. Media Psychology, Vol. 7(2), pp. 107-144 (2005).
18. Lazarus, R.S.: Emotion and Adaptation. New York: Oxford University (1991)
19. Lotka A. J.: Elements of Physical Biology. Reprinted by Dover in 1956 as Elements of Mathematical Biology (1924)
20. Marsella, S., and Gratch, J.: EMA: A Model of Emotional Dynamics. Cognitive Systems Research, Vol. 10(1), pp. 70-90 (2009).
21. Maynard S, J.: Models in Ecology. Cambridge University Press, Cambridge (1974).
22. Sabal, J. (2002). Financial Decisions in Emerging Markets. Oxford University Press, Inc., New York.
23. Thaler, R.H. "The Ultimatum Game." In: Journal of Economic Perspectives, volume 2, 1988, pp. 195-206.

24. Van Vugt, H.C., Hoorn, J.F., Konijn, E.A.: Interactive engagement with embodied agents, an empirically validated framework. Computer Animation and Virtual Worlds 20, 195-204 (2009).
25. Van Vugt, H.C., Konijn, E.A., Hoorn, J.F., Eliëns, A., & Keur, I. (2007). Realism is not all! User engagement with task-related interface characters. Interacting with Computers 19(2) 267-280.
26. Van Vugt, H.C., Konijn, E.A., Hoorn, J.F., & Veldhuis, J. (2006). Why fat interface characters are better e-health advisors. Lecture Notes in Artificial Intelligence (LNAI) 4133: 1-13. DOI 10.1007 / 11821830_1. Available at :http://www.springerlink.com /content/g379221v1w65tu38/fulltext.pdf.
27. Volterra, V.: Fluctuations in the abundance of a species considered mathematically. Nature vol. 118, pp. 558-560 (1926)
28. Worldviz Vizard Virtual Reality Toolkit. University of California. URL: http://www.worldviz.com/vizard.htm.

Samenvatting

Onderzoekers zijn de afgelopen jaren in toenemende mate geïnteresseerd geraakt in de toepassing van intelligente virtuele agenten in verschillende domeinen. Intelligente virtuele agenten (IVAs) zijn autonome, grafisch belichaamde agenten in een virtuele omgeving die in staat zijn om op een intelligente manier interactie te hebben met de omgeving, andere IVAs, of met menselijke gebruikers. Recentelijk is veel onderzoek gewijd aan het ontwikkelen van virtuele agenten met meer realistische grafische representaties. Echter, de affectieve eigenschappen van zulke agenten zijn doorgaans vrij beperkt, and niet erg menselijk. Hoewel veel IVAs tegenwoordig bijvoorbeeld de mogelijkheid hebben om op de één of andere manier emoties te tonen door middel van verschillende gezichtsuitdrukkingen, is het behoorlijk moeilijk voor ze om de juiste emotie op het juiste moment te tonen. Om nog een stap verder te gaan, het is zelfs nog moeilijker voor ze om de emotionele toestand van andere agenten daadwerkelijk te begrijpen en om daar op een empathische manier op te reageren. Dit is in tegenspraak met de behoefte van virtuele agenten om menselijk affectief gedrag nauwkeurig te kunnen nabootsen. Verschillende studies in de Sociale Wetenschappen hebben aangetoond dat dit een belangrijke voorwaarde is voor een agent om menselijke betrokkenheid binnen een virtuele omgeving te vergroten. Daarom zijn bestaande op IVAs gebaseerde systemen niet zo effectief als ze zouden kunnen zijn. Eigenschappen die zij typisch ontberen zijn de mogelijkheid om emoties te tonen (niet alleen in termen van gezichtsuitdrukking, maar ook in termen van gedrag), in samenhang met inzicht in de cognitieve en affectieve toestanden van elkaar en van mensen.

Om dit probleem aan te pakken stellen sommige auteurs voor om de affectieve eigenschappen van interactieve softwareagenten te vergroten door kennis van de Psychologie en Cognitiewetenschappen als basis te gebruiken voor het rekenkundig modelleren van de cognitieve en affectieve processen in kwestie. Recentelijk is een verscheidenheid van zulke computermodellen ontwikkeld voor verschillende aspecten van menselijk gedrag. Voorbeelden omvatten modellen voor redeneerprocessen, visuele aandacht, emotieregulatie, gedachten lezen, stress en werkdruk, en gemoedstoestanden. Wanneer zulke computermodellen beschikbaar zijn in een formele notatie, opent dit de mogelijkheid om IVAs met ze uit te rusten. Echter, de stap van bestaande computermodellen (die grotendeels gebruikt worden voor simulatiedoeleinden) naar modellen die direct kunnen worden ingeplugd in een virtuele (3D) omgeving, zodanig dat de IVAs zich volgens het cognitieve model gedragen, is een niet triviale. In werkelijkheid betreft deze stap een iteratief proces, bestaande uit, onder anderen, de volgende subtaken: het verfijnen van het computermodel, het vertalen naar een specifieke programmeeromgeving, en het testen en evalueren van het resulterende model in de virtuele setting.

Het voornaamste onderzoeksdoel van dit proefschrift is om te onderzoeken hoe computermodellen van emoties kunnen worden geïntegreerd binnen virtuele agenten. Hiertoe zijn verschillende theorieën vanuit verschillende gebieden (bv, Sociale Wetenschappen, Psychologie, Economie) beschouwd en gecombineerd om integratieve modellen van emoties te ontwikkelen. De meeste van deze modellen zijn eerst gebruikt voor simulatie, om te testen of hun gehele gedrag bevredigend was. Hiervoor zijn verschillende modelleeromgevingen gebruikt, zoals LEADSTO en C++. Vervolgens zijn deze modellen ingebouwd binnen toepassingen gerelateerd aan de zorg, games, en het bedrijfsleven. Deze toepassingen zijn ontwikkeld met behulp van JavaScript, in combinatie met de Vizard toolkit. Uiteindelijk zijn voor sommige toepassingen gebruikerstests uitgevoerd. De voorlopige resultaten van deze tests tonen aan dat gebruikers dankzij de ontwikkelde modellen meer betrokken raken bij de toepassingen.

In het eerste deel van dit proefschrift wordt de mate waarin een agent betrokken raakt bij een andere agent, of op afstand hiervan blijft, gemodelleerd, door op een rekenkundige manier het (eerder informeel beschreven) I-PEFiC model te formaliseren. Deze maten hangen af van verschillende aspecten van de andere agent zoals ethiek (goed of slecht), esthetiek (mooi of lelijk), realisme (hoe realistisch of onrealistisch de andere agent is), gelijkheid (overeenkomsten tussen de twee agenten) en mogelijkheden die de andere agent biedt als hulpmiddel of obstakel voor de taakuitvoering van de agent. Als tweede wordt de integratie van de drie modellen van emoties behandeld. De in dit deel gekozen aanpak is om drie van de meer invloedrijke modellen te selecteren, die gemeenschappelijk hebben dat ze kunnen worden gebruikt om geloofwaardigheid van virtuele karakters te vergroten: CoMERG (Computational Model of Emotion Regulation based on Gross theory), EMA (Emotion and Adaptation) en I-PEFiC^{ADM} (I-PEFiC uitgebreid met een module voor ‘Affective Decision Making’). Daarna focussen we op het modelleren van affectieve toestanden van een persoon in economische context. Deze economische context wordt beschouwd als een grootschalig multi-agentsysteem bestaande uit duizenden of miljoenen andere agenten. In dit deel wordt beschreven hoe iemands betrokkenheid bij deze andere agenten in de vorm van individuele investeringsbeslissingen afhangen van een persoonlijk risicoprofiel, de mate van hebzucht van de persoon, en de toestand van de wereldeconomie. Ten slotte worden in het laatste deel van het proefschrift toepassingen van virtuele agenten die emoties tonen in de domeinen van de zorg, het bedrijfsleven en games gepresenteerd en geëvalueerd.

SIKS Dissertation Series

2010

2010-01
Matthijs van Leeuwen (UU)
Patterns that Matter

2010-02
Ingo Wassink (UT)
Work flows in Life Science

2010-03
Joost Geurts (CWI)
A Document Engineering Model and Processing Framework for Multimedia documents

2010-04
Olga Kulyk (UT)
Do You Know What I Know? Situational Awareness of Co-located Teams in Multidisplay Environments

2010-05
Claudia Hauff (UT)
Predicting the Effectiveness of Queries and Retrieval Systems

2010-06
Sander Bakkes (UvT)
Rapid Adaptation of Video Game AI

2010-07
Wim Fikkert (UT)
A Gesture interaction at a Distance

2010-08
Krzysztof Siewicz (UL)
Towards an Improved Regulatory Framework of Free Software. Protecting user freedoms in a world of software communities and eGovernments

2010-09
Hugo Kielman (UL)
Politiële gegevensverwerking en Privacy, Naar een effectieve waarborging

2010-10
Rebecca Ong (UL)
Mobile Communication and Protection of Children

2010-11
Adriaan Ter Mors (TUD)
The world according to MARP: Multi-Agent Route Planning

2010-12
Susan van den Braak (UU)
Sensemaking software for crime analysis

2010-13
Gianluigi Folino (RUN)
High Performance Data Mining using Bio-inspired techniques

2010-14
Sander van Splunter (VU)
Automated Web Service Reconfiguration

2010-15
Lianne Bodenstaff (UT)
Managing Dependency Relations in Inter-Organizational Models

2010-16
Sicco Verwer (TUD)
Efficient Identification of Timed Automata, theory and practice

2010-17
Spyros Kotoulas (VU)
Scalable Discovery of Networked Resources: Algorithms, Infrastructure, Applications

2010-18
Charlotte Gerritsen (VU)
Caught in the Act: Investigating Crime by Agent-Based Simulation

2010-19
Henriette Cramer (UvA)
People's Responses to Autonomous and Adaptive Systems

2010-20
Ivo Swartjes (UT)
Whose Story Is It Anyway? How Improv Informs Agency and Authorship in Emergent Narrative

2010-21
Harold van Heerde (UT)
Privacy-aware data management by means of data degradation

2010-22
Michiel Hildebrand (CWI)
End-user Support for Access to\\ Heterogeneous Linked Data

2010-23
Bas Steunebrink (UU)
The Logical Structure of Emotions

2010-24
Dmytro Tykhonov
Designing Generic and Efficient Negotiation Strategies

2010-25
Zulfiqar Ali Memon (VU)
Modelling Human-Awareness for Ambient Agents: A Human Mindreading Perspective

2010-26
Ying Zhang (CWI)
XRPC: Efficient Distributed Query Processing on Heterogeneous XQuery Engines

2010-27
Marten Voulon (UL)
Automatisch contracteren

2010-28
Arne Koopman (UU)
Characteristic Relational Patterns

2010-29
Stratos Idreos (CWI)
Database Cracking: Towards Auto-tuning Database Kernels

2010-30
Marieke van Erp (UvT)
Accessing Natural History - Discoveries in data cleaning, structuring, and retrieval

2010-31
Victor de Boer (UVA)
Ontology Enrichment from Heterogeneous Sources on the Web

2010-32
Marcel Hiel (UvT)
An Adaptive Service Oriented Architecture: Automatically solving Interoperability Problems

2010-33
Robin Aly (UT)
Modeling Representation Uncertainty in Concept-Based Multimedia Retrieval

2010-34
Teduh Dirgahayu (UT)
Interaction Design in Service Compositions

2010-35
Dolf Trieschnigg (UT)
Proof of Concept: Concept-based Biomedical Information Retrieval

2010-36
José Janssen (OU)
Paving the Way for Lifelong Learning Facilitating competence development through a learning path specification

2010-37
Niels Lohmann (TUE)
Correctness of services and their composition

2010-38
Dirk Fahland (TUE)
From Scenarios to components

2010-39
Ghazanfar Farooq Siddiqui (VU)
Integrative Modeling of Emotions in Virtual Agents

2009

2009-01
Rasa Jurgelenaite (RUN)
Symmetric Causal Independence Models

2009-02
Willem Robert van Hage (VU)
Evaluating Ontology-Alignment Techniques

2009-03
Hans Stol (UvT)
A Framework for Evidence-based Policy Making Using IT

2009-04
Josephine Nabukenya (RUN)
Improving the Quality of Organisational Policy Making using Collaboration Engineering

2009-05
Sietse Overbeek (RUN)
Bridging Supply and Demand for Knowledge Intensive Tasks - Based on Knowledge, Cognition, and Quality

2009-06
Muhammad Subianto (UU)
Understanding Classification

2009-07
Ronald Poppe (UT)
Discriminative Vision-Based Recovery and Recognition of Human Motion

2009-08
Volker Nannen (VU)
Evolutionary Agent-Based Policy Analysis in Dynamic Environments

2009-09
Benjamin Kanagwa (RUN)
Design, Discovery and Construction of Service-oriented Systems

2009-10
Jan Wielemaker (UVA)
Logic programming for knowledge-intensive interactive applications

2009-11
Alexander Boer (UVA)
Legal Theory, Sources of Law & the Semantic Web

2009-12
Peter Massuthe (TUE, Humboldt-Universitaet zu Berlin)
Operating Guidelines for Services

2009-13
Steven de Jong (UM)
Fairness in Multi-Agent Systems

2009-14
Maksym Korotkiy (VU)
From ontology-enabled services to service-enabled ontologies (making ontologies work in e-science with ONTO-SOA)

2009-15
Rinke Hoekstra (UVA)
Ontology Representation - Design Patterns and Ontologies that Make Sense

2009-16
Fritz Reul (UvT)
New Architectures in Computer Chess

2009-17
Laurens van der Maaten (UvT)
Feature Extraction from Visual Data

2009-18
Fabian Groffen (CWI)
Armada, An Evolving Database System

2009-19
Valentin Robu (CWI)
Modeling Preferences, Strategic Reasoning and Collaboration in Agent-Mediated Electronic Markets

2009-20
Bob van der Vecht (UU)
Adjustable Autonomy: Controling Influences on Decision Making

2009-21
Stijn Vanderlooy (UM)
Ranking and Reliable Classification

2009-22
Pavel Serdyukov (UT)
Search For Expertise: Going beyond direct evidence

2009-23
Peter Hofgesang (VU)
Modelling Web Usage in a Changing Environment

2009-24
Annerieke Heuvelink (VU)
Cognitive Models for Training Simulations

2009-25
Alex van Ballegooij (CWI)
"RAM: Array Database Management through Relational Mapping"

2009-26
Fernando Koch (UU)
An Agent-Based Model for the Development of Intelligent Mobile Services

2009-27
Christian Glahn (OU)
Contextual Support of social Engagement and Reflection on the Web

2009-28
Sander Evers (UT)
Sensor Data Management with Probabilistic Models

2009-29
Stanislav Pokraev (UT)
Model-Driven Semantic Integration of Service-Oriented Applications

2009-30
Marcin Zukowski (CWI)
Balancing vectorized query execution with bandwidth-optimized storage

2009-31
Sofiya Katrenko (UVA)
A Closer Look at Learning Relations from Text

2009-32
Rik Farenhorst (VU) and Remco de Boer (VU)
Architectural Knowledge Management: Supporting Architects and Auditors

2009-33
Khiet Truong (UT)
How Does Real Affect Affect Affect Recognition In Speech?

2009-34
Inge van de Weerd (UU)
Advancing in Software Product Management: An Incremental Method Engineering Approach

2009-35
Wouter Koelewijn (UL)
Privacy en Politiegegevens; Over geautomatiseerde normatieve informatie-uitwisseling

2009-36
Marco Kalz (OU)
Placement Support for Learners in Learning Networks

2009-37
Hendrik Drachsler (OU)
Navigation Support for Learners in Informal Learning Networks

2009-38
Riina Vuorikari (OU)
Tags and self-organisation: a metadata ecology for learning resources in a multilingual context

2009-39
Christian Stahl (TUE, Humboldt-Universitaet zu Berlin)
Service Substitution -- A Behavioral Approach Based on Petri Nets

2009-40
Stephan Raaijmakers (UvT)
Multinomial Language Learning: Investigations into the Geometry of Language

2009-41
Igor Berezhnyy (UvT)
Digital Analysis of Paintings

2009-42
Toine Bogers (UvT)
Recommender Systems for Social Bookmarking

2009-43
Virginia Nunes Leal Franqueira (UT)
Finding Multi-step Attacks in Computer Networks using Heuristic Search and Mobile Ambients

2009-44
Roberto Santana Tapia (UT)
Assessing Business-IT Alignment in Networked Organizations

2009-45
Jilles Vreeken (UU)
Making Pattern Mining Useful

2009-46
Loredana Afanasiev (UvA)
Querying XML: Benchmarks and Recursion

2008

2008-01
Katalin Boer-Sorbán (EUR)
Agent-Based Simulation of Financial Markets: A modular, continuous-time approach

2008-02
Alexei Sharpanskykh (VU)
On Computer-Aided Methods for Modeling and Analysis of Organizations

2008-03
Vera Hollink (UVA)
Optimizing hierarchical menus: a usage-based approach

2008-04
Ander de Keijzer (UT)
Management of Uncertain Data - towards unattended integration

2008-05
Bela Mutschler (UT)
Modeling and simulating causal dependencies on process-aware information systems from a cost perspective

2008-06
Arjen Hommersom (RUN)
On the Application of Formal Methods to Clinical Guidelines, an Artificial Intelligence Perspective

2008-07
Peter van Rosmalen (OU)
Supporting the tutor in the design and support of adaptive e-learning

2008-08
Janneke Bolt (UU)
Bayesian Networks: Aspects of Approximate Inference

2008-09
Christof van Nimwegen (UU)
The paradox of the guided user: assistance can be counter-effective

2008-10
Wouter Bosma (UT)
Discourse oriented summarization

2008-11
Vera Kartseva (VU)
Designing Controls for Network Organizations: A Value-Based Approach

2008-12
Jozsef Farkas (RUN)
A Semiotically Oriented Cognitive Model of Knowledge Representation

2008-13
Caterina Carraciolo (UVA)
Topic Driven Access to Scientific Handbooks

2008-14
Arthur van Bunningen (UT)
Context-Aware Querying; Better Answers with Less Effort

2008-15
Martijn van Otterlo (UT)
The Logic of Adaptive Behavior: Knowledge Representation and Algorithms for the Markov Decision Process Framework in First-Order Domains.

2008-16
Henriette van Vugt (VU)
Embodied agents from a user's perspective

2008-17
Martin Op 't Land (TUD)
Applying Architecture and Ontology to the Splitting and Allying of Enterprises

2008-18
Guido de Croon (UM)
Adaptive Active Vision

2008-19
Henning Rode (UT)
From Document to Entity Retrieval: Improving Precision and Performance of Focused Text Search

2008-20
Rex Arendsen (UVA)
Geen bericht, goed bericht. Een onderzoek naar de effecten van de introductie van elektronisch berichtenverkeer met de overheid op de administratieve lasten van bedrijven.

2008-21
Krisztian Balog (UVA)
People Search in the Enterprise

2008-22
Henk Koning (UU)
Communication of IT-Architecture

2008-23
Stefan Visscher (UU)
Bayesian network models for the management of ventilator-associated pneumonia

2008-24
Zharko Aleksovski (VU)
Using background knowledge in ontology matching

2008-25
Geert Jonker (UU)
Efficient and Equitable Exchange in Air Traffic Management Plan Repair using Spender-signed Currency

2008-26
Marijn Huijbregts (UT)
Segmentation, Diarization and Speech Transcription: Surprise Data Unraveled

2008-27
Hubert Vogten (OU)
Design and Implementation Strategies for IMS Learning Design

2008-28
Ildiko Flesch (RUN)
On the Use of Independence Relations in Bayesian Networks

2008-29
Dennis Reidsma (UT)
Annotations and Subjective Machines - Of Annotators, Embodied Agents, Users, and Other Humans

2008-30
Wouter van Atteveldt (VU)
Semantic Network Analysis: Techniques for Extracting, Representing and Querying Media Content

2008-31
Loes Braun (UM)
Pro-Active Medical Information Retrieval

2008-32
Trung H. Bui (UT)
Toward Affective Dialogue Management using Partially Observable Markov Decision Processes

2008-33
Frank Terpstra (UVA)
Scientific Workflow Design; theoretical and practical issues

2008-34
Jeroen de Knijf (UU)
Studies in Frequent Tree Mining

2008-35
Ben Torben Nielsen (UvT)
Dendritic morphologies: function shapes structure

2007

2007-01
Kees Leune (UvT)
Access Control and Service-Oriented Architectures

2007-02
Wouter Teepe (RUG)
Reconciling Information Exchange and Confidentiality: A Formal Approach

2007-03
Peter Mika (VU)
Social Networks and the Semantic Web

2007-04
Jurriaan van Diggelen (UU)
Achieving Semantic Interoperability in Multi-agent Systems: a dialogue-based approach

2007-05
Bart Schermer (UL)
Software Agents, Surveillance, and the Right to Privacy: a Legislative Framework for Agent-enabled Surveillance

2007-06
Gilad Mishne (UVA)
Applied Text Analytics for Blogs

2007-07
Natasa Jovanovic' (UT)
To Whom It May Concern - Addressee Identification in Face-to-Face Meetings

2007-08
Mark Hoogendoorn (VU)
Modeling of Change in Multi-Agent Organizations

2007-09
David Mobach (VU)
Agent-Based Mediated Service Negotiation

2007-10
Huib Aldewereld (UU)
Autonomy vs. Conformity: an Institutional Perspective on Norms and Protocols

2007-11
Natalia Stash (TUE)
Incorporating Cognitive/Learning Styles in a General-Purpose Adaptive Hypermedia System

2007-12
Marcel van Gerven (RUN)
Bayesian Networks for Clinical Decision Support: A Rational Approach to Dynamic Decision-Making under Uncertainty

2007-13
Rutger Rienks (UT)
Meetings in Smart Environments; Implications of Progressing Technology

2007-14
Niek Bergboer (UM)
Context-Based Image Analysis

2007-15
Joyca Lacroix (UM)
NIM: a Situated Computational Memory Model

2007-16
Davide Grossi (UU)
Designing Invisible Handcuffs. Formal investigations in Institutions and Organizations for Multi-agent Systems

2007-17
Theodore Charitos (UU)
Reasoning with Dynamic Networks in Practice

2007-18
Bart Orriens (UvT)
On the development an management of adaptive business collaborations

2007-19
David Levy (UM)
Intimate relationships with artificial partners

2007-20
Slinger Jansen (UU)
Customer Configuration Updating in a Software Supply Network

2007-21
Karianne Vermaas (UU)
Fast diffusion and broadening use: A research on residential adoption and usage of broadband internet in the Netherlands between 2001 and 2005

2007-22
Zlatko Zlatev (UT)
Goal-oriented design of value and process models from patterns

2007-23
Peter Barna (TUE)
Specification of Application Logic in Web Information Systems

2007-24
Georgina Ramírez Camps (CWI)
Structural Features in XML Retrieval

2007-25
Joost Schalken (VU)
Empirical Investigations in Software Process Improvement

2006

2006-01
Samuil Angelov (TUE)
Foundations of B2B Electronic Contracting

2006-02
Cristina Chisalita (VU)
Contextual issues in the design and use of information technology in organizations

2006-03
Noor Christoph (UVA)
The role of metacognitive skills in learning to solve problems

2006-04
Marta Sabou (VU)
Building Web Service Ontologies

2006-05
Cees Pierik (UU)
Validation Techniques for Object-Oriented Proof Outlines

2006-06
Ziv Baida (VU)
Software-aided Service Bundling - Intelligent Methods & Tools for Graphical Service Modeling

2006-07
Marko Smiljanic (UT)
XML schema matching -- balancing efficiency and effectiveness by means of clustering

2006-08
Eelco Herder (UT)
Forward, Back and Home Again - Analyzing User Behavior on the Web

2006-09
Mohamed Wahdan (UM)
Automatic Formulation of the Auditor's Opinion

2006-10
Ronny Siebes (VU)
Semantic Routing in Peer-to-Peer Systems

2006-11
Joeri van Ruth (UT)
Flattening Queries over Nested Data Types

2006-12
Bert Bongers (VU)
Interactivation - Towards an e-cology of people, our technological environment, and the arts

2006-13
Henk-Jan Lebbink (UU)
Dialogue and Decision Games for Information Exchanging Agents

2006-14
Johan Hoorn (VU)
Software Requirements: Update, Upgrade, Redesign - towards a Theory of Requirements Change

2006-15
Rainer Malik (UU)
CONAN: Text Mining in the Biomedical Domain

2006-16
Carsten Riggelsen (UU)
Approximation Methods for Efficient Learning of Bayesian Networks

2006-17
Stacey Nagata (UU)
User Assistance for Multitasking with Interruptions on a Mobile Device

2006-18
Valentin Zhizhkun (UVA)
Graph transformation for Natural Language Processing

2006-19
Birna van Riemsdijk (UU)
Cognitive Agent Programming: A Semantic Approach

2006-20
Marina Velikova (UvT)
Monotone models for prediction in data mining

2006-21
Bas van Gils (RUN)
Aptness on the Web

2006-22
Paul de Vrieze (RUN)
Fundaments of Adaptive Personalisation

2006-23
Ion Juvina (UU)
Development of Cognitive Model for Navigating on the Web

2006-24
Laura Hollink (VU)
Semantic Annotation for Retrieval of Visual Resources

2006-25
Madalina Drugan (UU)
Conditional log-likelihood MDL and Evolutionary MCMC

2006-26
Vojkan Mihajlovic (UT)
Score Region Algebra: A Flexible Framework for Structured Information Retrieval

2006-27
Stefano Bocconi (CWI)
Vox Populi: generating video documentaries from semantically annotated media repositories

2006-28
Borkur Sigurbjornsson (UVA)
Focused Information Access using XML Element Retrieval

2005

2005-01
Floor Verdenius (UVA)
Methodological Aspects of Designing Induction-Based Applications

2005-02
Erik van der Werf (UM)
AI techniques for the game of Go

2005-03
Franc Grootjen (RUN)
A Pragmatic Approach to the Conceptualisation of Language

2005-04
Nirvana Meratnia (UT)
Towards Database Support for Moving Object data

2005-05
Gabriel Infante-Lopez (UVA)
Two-Level Probabilistic Grammars for Natural Language Parsing

2005-06
Pieter Spronck (UM)
Adaptive Game AI

2005-07
Flavius Frasincar (TUE)
Hypermedia Presentation Generation for Semantic Web Information Systems

2005-08
Richard Vdovjak (TUE)
A Model-driven Approach for Building Distributed Ontology-based Web Applications

2005-09
Jeen Broekstra (VU)
Storage, Querying and Inferencing for Semantic Web Languages

2005-10
Anders Bouwer (UVA)
Explaining Behaviour: Using Qualitative Simulation in Interactive Learning Environments

2005-11
Elth Ogston (VU)
Agent Based Matchmaking and Clustering - A Decentralized Approach to Search

2005-12
Csaba Boer (EUR)
Distributed Simulation in Industry

2005-13
Fred Hamburg (UL)
Een Computermodel voor het Ondersteunen van Euthanasiebeslissingen

2005-14
Borys Omelayenko (VU)
Web-Service configuration on the Semantic Web; Exploring how semantics meets pragmatics

2005-15
Tibor Bosse (VU)
Analysis of the Dynamics of Cognitive Processes

2005-16
Joris Graaumans (UU)
Usability of XML Query Languages

2005-17
Boris Shishkov (TUD)
Software Specification Based on Re-usable Business Components

2005-18
Danielle Sent (UU)
Test-selection strategies for probabilistic networks

2005-19
Michel van Dartel (UM)
Situated Representation

2005-20
Cristina Coteanu (UL)
Cyber Consumer Law, State of the Art and Perspectives

2005-21
Wijnand Derks (UT)
Improving Concurrency and Recovery in Database Systems by Exploiting Application Semantics

2004

2004-01
Virginia Dignum (UU)
A Model for Organizational Interaction: Based on Agents, Founded in Logic

2004-02
Lai Xu (UvT)
Monitoring Multi-party Contracts for E-business

2004-03
Perry Groot (VU)
A Theoretical and Empirical Analysis of Approximation in Symbolic Problem Solving

2004-04
Chris van Aart (UVA)
Organizational Principles for Multi-Agent Architectures

2004-05
Viara Popova (EUR)
Knowledge discovery and monotonicity

2004-06
Bart-Jan Hommes (TUD)
The Evaluation of Business Process Modeling Techniques

2004-07
Elise Boltjes (UM)
Voorbeeldig onderwijs; voorbeeldgestuurd onderwijs, een opstap naar abstract denken, vooral voor meisjes

2004-08
Joop Verbeek (UM)
Politie en de Nieuwe Internationale Informatiemarkt, Grensregionale politiële gegevensuitwisseling en digitale expertise

2004-09
Martin Caminada (VU)
For the Sake of the Argument; explorations into argument-based reasoning

2004-10
Suzanne Kabel (UVA)
Knowledge-rich indexing of learning-objects

2004-11
Michel Klein (VU)
Change Management for Distributed Ontologies

2004-12
The Duy Bui (UT)
Creating emotions and facial expressions for embodied agents

2004-13
Wojciech Jamroga (UT)
Using Multiple Models of Reality: On Agents who Know how to Play

2004-14
Paul Harrenstein (UU)
Logic in Conflict. Logical Explorations in Strategic Equilibrium

2004-15
Arno Knobbe (UU)
Multi-Relational Data Mining

2004-16
Federico Divina (VU)
Hybrid Genetic Relational Search for Inductive Learning

2004-17
Mark Winands (UM)
Informed Search in Complex Games

2004-18
Vania Bessa Machado (UvA)
Supporting the Construction of Qualitative Knowledge Models

2004-19
Thijs Westerveld (UT)
Using generative probabilistic models for multimedia retrieval

2004-20
Madelon Evers (Nyenrode)
Learning from Design: facilitating multidisciplinary design teams

2003

2003-01
Heiner Stuckenschmidt (VU)
Ontology-Based Information Sharing in Weakly Structured Environments

2003-02
Jan Broersen (VU)
Modal Action Logics for Reasoning About Reactive Systems

2003-03
Martijn Schuemie (TUD)
Human-Computer Interaction and Presence in Virtual Reality Exposure Therapy

2003-04
Milan Petkovic (UT)
Content-Based Video Retrieval Supported by Database Technology

2003-05
Jos Lehmann (UVA)
Causation in Artificial Intelligence and Law - A modelling approach

2003-06
Boris van Schooten (UT)
Development and specification of virtual environments

2003-07
Machiel Jansen (UvA)
Formal Explorations of Knowledge Intensive Tasks

2003-08
Yongping Ran (UM)
Repair Based Scheduling

2003-09
Rens Kortmann (UM)
The resolution of visually guided behaviour

2003-10
Andreas Lincke (UvT)
Electronic Business Negotiation: Some experimental studies on the interaction between medium, innovation context and culture

2003-11
Simon Keizer (UT)
Reasoning under Uncertainty in Natural Language Dialogue using Bayesian Networks

2003-12
Roeland Ordelman (UT)
Dutch speech recognition in multimedia information retrieval

2003-13
Jeroen Donkers (UM)
Nosce Hostem - Searching with Opponent Models

2003-14
Stijn Hoppenbrouwers (KUN)
Freezing Language: Conceptualisation Processes across ICT-Supported Organisations

2003-15
Mathijs de Weerdt (TUD)
Plan Merging in Multi-Agent Systems

2003-16
Menzo Windhouwer (CWI)
Feature Grammar Systems - Incremental Maintenance of Indexes to Digital Media Warehouses

2003-17
David Jansen (UT)
Extensions of Statecharts with Probability, Time, and Stochastic Timing

2003-18
Levente Kocsis (UM)
Learning Search Decisions

2002

2002-01
Nico Lassing (VU)
Architecture-Level Modifiability Analysis

2002-02
Roelof van Zwol (UT)
Modelling and searching web-based document collections

2002-03
Henk Ernst Blok (UT)
Database Optimization Aspects for Information Retrieval

2002-04
Juan Roberto Castelo Valdueza (UU)
The Discrete Acyclic Digraph Markov Model in Data Mining

2002-05
Radu Serban (VU)
The Private Cyberspace Modeling Electronic Environments inhabited by Privacy-concerned Agents

2002-06
Laurens Mommers (UL)
Applied legal epistemology; Building a knowledge-based ontology of the legal domain

2002-07
Peter Boncz (CWI)
Monet: A Next-Generation DBMS Kernel For Query-Intensive Applications

2002-08
Jaap Gordijn (VU)
Value Based Requirements Engineering: Exploring Innovative E-Commerce Ideas

2002-09
Willem-Jan van den Heuvel (KUB)
Integrating Modern Business Applications with Objectified Legacy Systems

2002-10
Brian Sheppard (UM)
Towards Perfect Play of Scrabble

2002-11
Wouter C.A. Wijngaards (VU)
Agent Based Modelling of Dynamics: Biological and Organisational Applications

2002-12
Albrecht Schmidt (Uva)
Processing XML in Database Systems

2002-13
Hongjing Wu (TUE)
A Reference Architecture for Adaptive Hypermedia Applications

2002-14
Wieke de Vries (UU)
Agent Interaction: Abstract Approaches to Modelling, Programming and Verifying Multi-Agent Systems

2002-15
Rik Eshuis (UT)
Semantics and Verification of UML Activity Diagrams for Workflow Modelling

2002-16
Pieter van Langen (VU)
The Anatomy of Design: Foundations, Models and Applications

2002-17
Stefan Manegold (UVA)
Understanding, Modeling, and Improving Main-Memory Database Performance

2001

2001-1
Silja Renooij (UU)
Qualitative Approaches to Quantifying Probabilistic Networks

2001-2
Koen Hindriks (UU)
Agent Programming Languages: Programming with Mental Models

2001-3
Maarten van Someren (UvA)
Learning as problem solving

2001-4
Evgueni Smirnov (UM)
Conjunctive and Disjunctive Version Spaces with Instance-Based Boundary Sets

2001-5
Jacco van Ossenbruggen (VU)
Processing Structured Hypermedia: A Matter of Style

2001-6
Martijn van Welie (VU)
Task-based User Interface Design

2001-7
Bastiaan Schonhage (VU)
Diva: Architectural Perspectives on Information Visualization

2001-8
Pascal van Eck (VU)
A Compositional Semantic Structure for Multi-Agent Systems Dynamics

2001-9
Pieter Jan 't Hoen (RUL)
Towards Distributed Development of Large Object-Oriented Models, Views of Packages as Classes

2001-10
Maarten Sierhuis (UvA)
Modeling and Simulating Work Practice BRAHMS: a multiagent modeling and simulation language for work practice analysis and design

2001-11
Tom M. van Engers (VUA)
Knowledge Management: The Role of Mental Models in Business Systems Design

2000

2000-1
Frank Niessink (VU)
Perspectives on Improving Software Maintenance

2000-2
Koen Holtman (TUE)
Prototyping of CMS Storage Management

2000-3
Carolien M.T. Metselaar (UVA)
Sociaal-organisatorische gevolgen van kennistechnologie; een procesbenadering en actorperspectief.

2000-4
Geert de Haan (VU)
ETAG, A Formal Model of Competence Knowledge for User Interface Design

2000-5
Ruud van der Pol (UM)
Knowledge-based Query Formulation in Information Retrieval.

2000-6
Rogier van Eijk (UU)
Programming Languages for Agent Communication

2000-7
Niels Peek (UU)
Decision-theoretic Planning of Clinical Patient Management

2000-8
Veerle Coupé (EUR)
Sensitivity Analyis of Decision-Theoretic Networks

2000-9
Florian Waas (CWI)
Principles of Probabilistic Query Optimization

2000-10
Niels Nes (CWI)
Image Database Management System Design Considerations, Algorithms and Architecture

2000-11
Jonas Karlsson (CWI)
Scalable Distributed Data Structures for Database Management

1999

1999-1
Mark Sloof (VU)
Physiology of Quality Change Modelling; Automated modelling of Quality Change of Agricultural Products

1999-2
Rob Potharst (EUR)
Classification using decision trees and neural nets

1999-3
Don Beal (UM)
The Nature of Minimax Search

1999-4
Jacques Penders (UM)
The practical Art of Moving Physical Objects

1999-5
Aldo de Moor (KUB)
Empowering Communities: A Method for the Legitimate User-Driven Specification of Network Information Systems

1999-6
Niek J.E. Wijngaards (VU)
Re-design of compositional systems

1999-7
David Spelt (UT)
Verification support for object database design

1999-8
Jacques H.J. Lenting (UM)
Informed Gambling: Conception and Analysis of a Multi-Agent Mechanism for Discrete Reallocation.

1998

1998-1
Johan van den Akker (CWI)
DEGAS - An Active, Temporal Database of Autonomous Objects

1998-2
Floris Wiesman (UM)
Information Retrieval by Graphically Browsing Meta-Information

1998-3
Ans Steuten (TUD)
A Contribution to the Linguistic Analysis of Business Conversations within the Language/Action Perspective

1998-4
Dennis Breuker (UM)
Memory versus Search in Games

1998-5
E.W.Oskamp (RUL)
Computerondersteuning bij Straftoemeting